May 2020

Three Essays on the Resiliency of Urban Infrastructure and Public Systems

Caitlin O'Loughlin
Clemson University, cnt@g.clemson.edu

Three Essays on the Resiliency of Urban Infrastructure and Public Systems

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Economics

by
Caitlin O'Loughlin
May 2020

Accepted by:
Dr. Paul W. Wilson, Committee Chair
Dr. Aleda Roth
Dr. Frederick Hanssen
Dr. Kevin Tsui

Abstract

My dissertation focuses on issues related to the resiliency of urban infrastructure and public systems. In the first chapter, I examine the effect of the financial condition of local governments on housing values and depreciation of the U.S. housing stock. Housing values bear the burden of municipal fiscal stress reflecting prospective and current homeowners reduced willingness-to-pay for housing. Using panel data from the American Housing Survey from 1984 to 2011, I estimate linear, quantile, and semiparametric varying coefficient (VC) models to examine these effects. The findings from the linear and quantile estimation are compared to the estimation of the VC models, which allow for a nonparametric, smoothed specification of building age.

The results suggest that aspects of municipal solvency have differential effects across the distribution of housing values and building age. New, lower priced houses see the largest increases in housing values from larger cash and long-run solvency ratios, which reflect greater spending on infrastructure and other long-run capital investment projects, whereas older, lower priced houses benefit from increases in service-level solvency, suggestive of greater spending on public amenities. Moreover, the results indicate that the housing stock depreciates more slowly in municipalities with larger values of revenue and expenditure per capita, with implied annual depreciation rates ranging from over 0.4 to 0.7 percent.

The second chapter examines the performance of U.S. municipal governments prior to, during, and following the financial crisis over the period 1997–2012. Fully nonparametric methods are employed to estimate technical efficiencies of cities, both over time and by U.S. Census Region, utilizing recently-developed statistical results. The results strongly suggest non-convexity of the local governments' production set, calling into question the results of previous studies examining municipal efficiency. Furthermore, the results suggest that mean efficiency and productivity declined in several U.S. regions during the financial crisis, and in some instances have not returned to their pre-crisis level.

Finally, in the third chapter, I look at the impact of a regulation change in San Francisco restricting owner move-in (OMI) evictions. In the 1990s, San Francisco saw a large number of OMI evictions, a no-fault eviction frequently used by landlords to remove units from the rent-controlled market. To prevent landlords from further exploiting the use of OMI evictions, San Francisco passed Proposition G in 1998. This regulation imposed higher costs on landlords utilizing this eviction type by restricting the usage of OMI evictions in San Francisco rent-controlled buildings. My identification strategy is supported by the use of a stochastic rent frontier which allows me to obtain a measure of housing quality that is a linear function of observables, thereby capturing the causal impact of the regulation on housing quality and allowing me to observe trends in tenant filtering.

Using household-level renter data from the American Housing Survey, my results indicate that following Proposition G, rent-controlled units in San Francisco saw lower housing quality. Additionally, I find that this result holds even after controlling for the presence of an on-site building owner. These findings are likely driven by lower levels of building maintenance and upkeep, reflecting the higher costs resulting from Proposition G. These findings suggest that while this regulation intended to preserve the stock of controlled housing in San Francisco, it resulted in several unintended consequences.

Dedication

To my husband Michael, my dog Whitman, and my son. This is for you guys.

Acknowledgments

I would like to thank my advisor, Dr. Paul W. Wilson (Paul). He has kept me focused and on track throughout my graduate school experience. It's been a pleasure to get to know you and an honor to work with you. Thank you sincerely for everything.

I would also like to express sincere appreciation to my committee members, Dr. Aleda Roth, Dr. Kevin Tsui, and Dr. Frederick Hanssen for your helpful comments and suggestions over the years.

I am extremely grateful to Michael and my family and friends who have supported me endlessly during this journey. I appreciate all the love and encouragement.

A huge thank you to all the NSF NRT faculty, students, and staff who have pushed me out of my comfort zone and introduced me to different areas of research and thought. You all have broadened my horizon and interests!

Finally, to my office mates of Sirrine 417, thank you for keeping me motivated. To Sarah and Natalia especially, who I could not have done this without.

Table of Contents

List of Tables

List of Figures

Chapter 1

The Capitalization of the "Shadow Mortgage" and its Implication for Housing Values and Depreciation

1.1 Introduction

Investment in real estate is critical to local governments as many rely on property taxes as a main source of revenue. It is in the best interest of both local governments and homeowners to promote and sustain a thriving real estate market. Since housing is both a consumption and an investment good, households may purchase real estate for different reasons. When deciding to buy, the decision of where to purchase or invest in real estate can be an important one. One may consider the size and attributes of the unit, as well as its location, which reflect the neighborhood characteristics and amenities of the surrounding area such as school quality, access to parks and recreation, and public infrastructure. While the characteristics of the home may be readily apparent, it may be more difficult to assess the financial condition of a city or county. Municipal financial health directly affects the provision of amenities and ultimately impact one's investment over time.

Using a traditional hedonic approach, the willingness-to-pay for certain location amenities such as environmental quality and school quality has been investigated in prior studies. I assert that the financial state of a municipality directly impacts the provision and quality of public goods and services. When a municipality is under fiscal stress, it may decrease the provision or quality of these amenities, resulting in reduced willingness-to-pay for housing. This fiscal stress may also impact the rate of depreciation of the housing stock, thereby eroding households' returns on investment, an issue which has not yet been addressed in the literature. These issues are of increasing importance with numerous municipalities filing for Chapter 9 bankruptcy in recent years including Vallejo, California (CA) in 2008; Harrisburg, Pennsylvania in 2011; Jefferson County, Alabama in 2011; Stockton, CA in 2012; San Bernardino, CA in 2012; Detroit, Michigan in 2013; and Hillview, Kentucky in 2015.

In this paper I investigate the capitalization of municipal fiscal stress in housing values and its impact on depreciation. This capitalization amounts to an additional shadow mortgage on housing. Using panel data from 1984 to 2011, I estimate linear and quantile regression hedonic pricing models, as well as semiparametric varying coefficient (VC) models, for a large and representative sample of owner and renter-occupied properties. While there are different approaches in the literature as to how to predict fiscal stress, I use widely accepted measures of municipal solvency including cash, budgetary, service-level, and long-run.

My findings provide evidence for the shadow mortgage and highlight the differential impact of municipal solvency aspects on housing values. Notably, increases in both cash and long-run solvency measures, suggesting a less financially stable government, negatively impact low-value properties. I find that a ten percent increase in total expenditure per capita leads to an approximate 0.6 percent increase in housing values at the 0.1 quantile, around $400. By contrast, at the 0.9 quantile, this results in almost a one percent decrease, over $3,500. Unsurprisingly, higher taxes per capita negatively impact homeowners across the distribution. The ability of a local government to balance its budget is positively valued by homeowners across the distribution of housing values. At the 0.5 quantile, a ten percent increase in total revenue to total expenditure results in a $1,400 increase in housing values. After accounting for different aspects of municipal solvency, I find that depreciation is lower in areas with higher levels of per capita spending on amenities compared to areas with larger cash or other long-run solvency measures.

The decision to rent or buy a home likely depends on many factors including household income and preferences. Henderson and Ioannides (1986) find that consumers tend to smooth consumption over time, suggesting households that expect lower income in the future (e.g. retirement) are more likely to purchase a home rather than rent. In addition to considering attributes of the home such as number of rooms and floor area, prospective homeowners may also value things such as low crime rates or high school quality. Pope and Pope (2012) investigate the impact of crime on housing values and find that it has a significant and negative effect, reflecting homeowners' willingness-to-avoid. Brasington and Hite (2005) estimate the demand for environmental quality and show that environmental quality and school quality are complement goods, as indicated by a positive cross-elasticity, whereas environmental quality and house size are substitutes. Current and prospective homeowners are responsive to changes in school quality. Hayes and Taylor (1996) find that while location is usually the main driver in residential housing prices, school quality also plays a significant role. Moreover, findings from Thompson (2016) show a decline in housing prices in response to a school district being labeled as fiscally stressed.

Local governments may take on debt by issuing bonds to fund public goods and services such as public works projects or other capital outlays. Bronshtein (2017) finds that governments that expect an increase in the local housing tax base may be more likely to take on debt which was evident in the last recession. This can have negative implications for homeowners with results from MacKay (2014) suggesting an overcapitalization of fiscal debt in home prices. Contributing to municipal debt concerns are rising unfunded pension obligations. Arnott and Meulbroek (2018) discuss the potential impacts of unfunded municipal pension debt and assert that since homes are fixed assets, homeowners will bear the full burden of these obligations over time through lower housing prices. Brinkman et al. (2016) examine determinants of municipal pension funding and find that pension funding choices are fully capitalized in land prices.

In addition to housing and location characteristics, prospective and current homeowners may consider the depreciation rate of housing since it is likely to impact their investment over time. Hulten and Wykoff (1980) define economic depreciation as the "decline in asset price (or shadow price) due to aging." Malpezzi et al. (1987) cite this definition in their estimation of depreciation at the metropolitan statistical area (MSA) level. Leigh (1980) estimates depreciation of the national housing stock to be between 0.2 and 0.4 percent annually. Palmquist (1979) finds that owner-occupied residences depreciate at around 0.8 percent annually, while Randolph (1988) estimates annual depreciation of 0.6 percent for renter-occupied properties, although he notes that it is not possible to identify depreciation without assuming away vintage effects, or obsolescence of the housing unit over time. Knight and Sirmans (1996) and Wilhelmsson (2008) show that houses with lower levels of maintenance expenditures tend to depreciate faster than those that are relatively more maintained. Francke and van de Minne (2017) find that after 50 years of no maintenance, a structure loses 43 percent of its value on average. For single-family homes, results from Harding et al. (2007) suggest homeowners see little economic gain after controlling for maintenance and home improvements.

Walters (2009) examines whether depreciation rates differ between subsidized and unsubsidized units. He expects subsidized units to depreciate more quickly since their rent only depends on meeting minimum quality standards, but finds no evidence to suggest a significant difference between the two groups. Galbraith (1998) examines filtering as a low-income housing policy and notes that while overall consumption of housing has increased over time, the changes were larger for relatively wealthier households. The location choice of low and high-income households may help to explain this finding. It could be that wealthier households are attracted to amenities, both built and natural, in the city center. Brueckner et al. (1999) argue that this could lead to a Paris-style location pattern with wealthier households concentrating in the city center and households with relatively less means populating the outer arrondissements. Alternatively, the age of the housing stock may explain where high and low-income households choose to live. Brueckner and Rosenthal (2009) find that high-income households tend to locate in areas with a relatively young housing stock. These findings lend support to the idea of downward filtering, or housing trickling down from the wealthiest households to the poorest, until the lowest quality housing drops out of the stock.

The impact of the shadow mortgage on housing values and investment motivates the research question. To my knowledge, no previous studies have empirically analyzed this question using the framework and methodology proposed here. Moreover, the impact of local government fiscal stress on the housing stock is increasingly important in the post-recession climate. This paper expands on the existing literature by examining the differential effect of aspects of the municipal financial condition on housing values and also provides a framework to account for its impact on depreciation.

The rest of the paper is organized as follows. In Section 1.2, I present a simple theoretical model of bid rent and housing value. I develop the statistical model in Section 1.3. In Section 1.4, I describe the data and provide summary statistics. The empirical findings are presented in Section 1.5. In Section 1.6, I conclude and summarize the results of the paper as well as describe additional findings and robustness checks.

1.2 Theoretical Model

I start with a simple theoretical model, adapting from the framework of Rosen (1974) and Hayes and Taylor (1996). I assume that consumers are rational and attempt to maximize their utility, taking the housing stock as given. Consumers earn income y and derive utility from consumption of z, a vector of housing characteristics with prices p, and x, a composite good. Consumers attempt to maximize utility

$$U = U(x, z_1, z_2, \ldots, z_n) \tag{2.1}$$

subject to their budget constraint

$$y = x + p_1 z_1 + p_2 z_2 + \ldots + p_n z_n, \tag{2.2}$$

where x is the numeraire good. Therefore, the consumer maximization problem is

$$\mathcal{L} = U(x, z_1, z_2, \ldots, z_n) + \lambda(y - x - p_1 z_1 - p_2 z_2 - \ldots - p_n z_n) \tag{2.3}$$

with first-order conditions

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial U}{\partial x} - \lambda = 0 \tag{2.4}$$

and

$$\frac{\partial \mathcal{L}}{\partial z_i} = \frac{\partial U}{\partial z_i} - \lambda p_i = 0 \tag{2.5}$$

for the i^{th} housing characteristic, $i = 1, \ldots, n$. From (2.4) and (2.5), it can be shown that the marginal rate of substitution between good x and z_i is equal to the ratio of prices

$$\frac{U_x}{U_{z_i}} = p_i. \tag{2.6}$$

The demand equations for x^* and z_i^*, $i = 1, \ldots, n$, can be obtained from (2.6). Following Henderson (1977), after substituting the demand equations into the utility function, the consumer's indirect utility function is

$$U^*(y - R, z_1, z_2, \ldots, z_n) \tag{2.7}$$

where R represents total expenditures on housing services. The consumer's bid rent function is therefore consumer's willingness-to-pay for values of z at a given level of U^* and y. After taking the inverse of (2.7), it can be shown that the consumer's bid rent function is

$$R = R(z_1, z_2, \ldots, z_n \mid y, U^*) \tag{2.8}$$

where U^* is the level of indirect utility from (2.7). To estimate a consumer's willingness-to-pay for certain characteristics, partial derivatives of the bid rent function with respect to characteristic z_i can be taken.

The present value of housing services, V, to a potential homebuyer is the discounted sum of after-tax bid rents. Letting τ be the tax rate, i the discount rate, and assuming that housing is an infinitely lived asset, then

$$V = \sum_{t=0}^{\infty}(R - \tau V)e^{-rt} = \frac{R - \tau V}{i} \tag{2.9}$$

or equivalently,

$$V = \frac{R}{i + \tau}. \tag{2.10}$$

In equilibrium, the value of housing services equals the highest bid offered by potential consumers. To model the value of housing services, a hedonic pricing function can be estimated containing all characteristics of housing services conferred including observable unit attributes, as well as neighborhood and location amenities such as high-quality schools, highways, and public parks. I posit the revenues and expenditures of municipalities directly fund these public goods and services and the financial state can impact the provision and quality of these goods.

1.3 Statistical Model

The value of housing services is a function of housing and location specific characteristics. To estimate the marginal effect of each attribute on the response variable, I construct a hedonic model consistent with the proposed theoretical model. Since there is not a renter-equivalent measure of housing value in the data, I specify and estimate a hedonic pricing model for the owner and renter samples separately. My identification strategy consists of two key assumptions: the existence of a national housing market and that the municipal solvency measures accurately reflect the amount and quality of publicly provided amenities. The first assumption is supported by Linneman (1980) who finds that tests of the national housing market hypotheses cannot be rejected. Previous studies have utilized this assumption including

Chay and Greenstone (2005) who examine the capitalization of air quality in housing values. Identification of the municipal financial state comes through a broad set of observations, with a wide range of solvency measures and fiscal conditions.

Typically, the coefficient on the building age term of the hedonic pricing function is used to estimate depreciation, although this may be misleading if there are nonlinear or interaction effects. To account for these possibilities, I specify an age function

$$h(A_{it}, D_i, F_{it}, S_{ct}) = \sum_{k=1}^{10} \gamma_k a_{it} \mathbb{I}_k(\alpha_{k-1} < a_{it} \leq \alpha_k) + \gamma_{11} a_{it} D_i + \gamma_{12} a_{it} F_{it} + \gamma_{13} a_{it} S_{ct}, \tag{3.11}$$

where A_{it} is a vector of interactions between the building age of household i at time t, a_{it}, and building age group indicators $\mathbb{I}_k$ following Yoshida (2016), such that α_{k-1} and α_k represent the lower and upper bounds on each age group, respectively. The vector S_{ct} includes the solvency measures for county c at time t, D_i is an indicator for detached housing units, and F_{it} is the floor area in square feet. The age function specification allows depreciation to vary over the life of the structure and also with structure type, floor area, and the municipal financial condition.

To capture the age profile of housing and any potential nonlinear effects, I discretize the first building age term appearing in (3.11). I include an indicator for detached structures since it is likely that they depreciate more slowly than other types of structures. Randolph (1988) notes that the building age interaction with structure type is especially important because the economics and technology of maintenance behavior is likely to vary across building types. Similarly, depreciation is likely to vary with house size, or floor area, since larger structures tend to be located farther away from the city center where land prices are relatively cheaper. Finally, the interaction between building age and the solvency measure is justified by Breger (1967) who states that depreciation of property may result from either deterioration of the capacity to render service or a decline in the demand for the service rendered.

For the owner-occupied specification, the complete reduced-form hedonic pricing model is

$$lnV_{it} = \alpha_0 + h(A_{it}, D_i, F_{it}, S_{ct}) + X_{it}\beta_1 + Z_{it}\beta_2 + S_{ct}\beta_3 + \epsilon_{it}, \tag{3.12}$$

where lnV_{it} is the log of the self-reported housing value for household i at time t. While it is possible that homeowners may not precisely report the value of their home, results from Robins and West (1977) reveal precision between homeowners' self-reported and appraised values. Kiel and Zabel (1999) further find that self-reported measures tend to yield reliable estimates, noting that there is no systematic evidence to suggest certain groups of homeowners more accurately assess home value than others. The vector X_{it} contains observable unit characteristics and household demographics such as the number of rooms, bathrooms, and household size, Z_{it} contains community controls and geographic information including the zone-level average household income and education, and S_{ct} includes the municipal financial state of county c at time t as defined by the cash, budgetary, service, and long-run solvency measures. Since the dependent variable is an individual measure of housing value, as opposed to an aggregate measure, concerns over simultaneity can be avoided.

All solvency measures are for the prior fiscal year since I expect there to be a lag effect in terms of the provision of amenities and services. In this model specification, the marginal effect of the solvency measure on the response variable is a linear function of building age. To obtain a measure of depreciation, I estimate the marginal effect

$$\frac{\partial ln(V_{it})}{\partial h(A_{it}, D_i, F_{it}, S_{ct})} \frac{\partial h(A_{it}, D_i, F_{it}, S_{ct})}{\partial a_{it}}, \tag{3.13}$$

of building age on housing values evaluated at different percentiles of the solvency measures. As homes age, their value typically decreases over time, although Chinloy (1978) notes that theoretically there is no reason for the depreciation function to be downward sloping.

In addition to the estimation of (3.12), a similar model, i.e.,

$$lnR_{it} = \gamma_0 + g(A_{it}, D_i, F_{it}, S_{ct}) + X^*_{it}\delta_1 + Z_{it}\delta_2 + S_{ct}\delta_3 + \varepsilon_{it}, \tag{3.14}$$

is estimated for the renter-occupied sample where lnR_{it} is the log of the annual contract rent of household i at time t. The age function $g(A_{it}, D_i, F_{it}, S_{ct})$ is similar to the one for owner-occupiers in (3.11). The baseline model for both the owner and renter sample includes the full set of controls, age function, as well as location and year fixed effects, but excludes any of the municipal solvency terms or interactions with these measures.[1]

I estimate (3.12) and (3.14) using both ordinary least squares (OLS) and quantile regression. The estimation of the hedonic pricing model via OLS can be severely distorted by outlier observations making quantile regression an attractive approach. Furthermore, it is likely that the marginal effect of municipal solvency differs across the distribution of housing values. Zietz et al. (2008) use quantile regression to estimate a hedonic pricing model to examine the willingness-to-pay for certain housing characteristics. They find that owners of relatively higher-priced homes value certain characteristics such as square footage and number of bathrooms differently than lower priced ones. They also note that the distribution of age also varies across the different quantiles examined. Zahirovic-Herbert and Chatterjee (2012) find that his-

[1] It should be noted that annual property taxes and maintenance costs are omitted from the renter sample estimation since there is no renter equivalent in the data. This is reflected in X^*_{it} and is discussed in section 1.4.

toric preservation tends to positively impact relatively low-price houses, but the act of preservation may displace low-income residents, leading to a quasi-gentrification effect. By contrast, Zhang (2016) finds that lower valued homes bear more of the negative impact following a major flood.

The quantile regression model for the owner sample, Q, due to Koenker and following Greene (2000) can be written as

$$Q(\mathrm{V} \mid \mathbf{Z}, q) = \mathbf{Z}'\beta_{\mathbf{q}} \text{ such that } \Pr(\mathrm{V} \leq \mathbf{Z}'\beta_{\mathbf{q}} \mid \mathbf{Z}) = q,\ 0 < q < 1, \tag{3.15}$$

where V represents the log of housing value, vector $\mathbf{Z}$ includes the right-hand side of (3.12) and q represents a given quantile strictly between 0 and 1. Since no assumption is made about the distribution of $\mathrm{V} \mid \mathbf{Z}$ or about its conditional variance, this is essentially a semiparametric specification. The estimator, $\mathbf{b}_q$ of β_q for a specific quantile is computed by minimizing the function

$$\begin{aligned} F_n(\beta_q \mid \mathbf{V}, \mathbf{Z}) &= \sum_{i:V_i \geq Z_i'\beta_q}^{n} q \mid V_i - \mathbf{Z_i}'\beta_{\mathbf{q}} \mid + \sum_{i:V_i < Z_i'\beta_q}^{n} (1-q) \mid V_i - \mathbf{Z_i}'\beta_{\mathbf{q}} \mid \\ &= \sum_{i=1}^{n} g(V_i - \mathbf{Z}_i'\beta_q \mid q) \end{aligned}$$

$$g(e_{i,q} \mid q) = \begin{cases} qe_{i,q} & \text{if } e_{i,q} \geq 0 \\ (1-q)e_{i,q} & \text{if } e_{i,q} < 0, \end{cases} \tag{3.16}$$

and $e_{i,q} = V_i - \mathbf{Z}_i'\beta_q$.[2] To make inference, I bootstrap and cluster the standard errors at the county-level. All models are estimated at the 0.1, 0.25, 0.5, 0.75, and 0.9

[2]The quantile regression model for the owner-occupied sample in (3.15)–(3.16) can be easily be modified for the renter sample. The separate appendix can be requested for results from the renter sample estimation.

quantiles of housing values.

Since I expect the effect of the solvency measures on housing values to vary with building age, I estimate a semiparametric VC model due to Hastie and Tibshirani (1993) as a robustness check for my hedonic pricing models. It is similar to a general additive model in that a single term enters the function nonparametrically, although VC models permit an interaction between nonparametric and parametric terms. This approach offers several advantages over a standard parametric framework. Specifically, these models relax the assumption of linearity between the predictors and response variables. Wan et al. (2017) apply the VC modeling framework to the estimation of hedonic house prices functions in Hong Kong. They note that VC models reduce modeling bias and also avoid the curse of dimensionality, both advantages over the traditional parametric model.

The VC model specification replaces the parametric building age terms in (3.11) with smoothed nonparametric versions. The modified building age function for the owner-occupied sample can be represented as

$$\widetilde{h}(a_{it}, D_i, F_{it}, S_{ct}) = f_1(a_{it}) + f_2(a_{it})D_i + f_3(a_{it})F_{it} + f_4(a_{it})S_{ct}, \tag{3.17}$$

where $f_j(a_{it})$ are the smooth building age functions, $j = 1, \ldots, 4$. The modified building age function $\widetilde{g}(a_{it}, D_i, F_{it}, S_{ct})$ for the renter sample is similar to (3.17). I adapt (3.12) and (3.14) with the modified building age functions for the owner and renter sample estimation, respectively.

1.4 Data

Data used for estimation are obtained from the American Housing Survey (AHS), U.S. Census of Governments, U.S. Census Bureau, and Bureau of Labor Statistics (BLS). The primary source of data is the AHS. The AHS includes both a national survey and metropolitan survey, with observations in the metropolitan sample being tracked over time. I use the metropolitan sample data which allow me to identify the household county. Unfortunately, not every county is sampled in every survey and therefore the number of years between observations is nonconstant. AHS data include characteristics of the housing unit as well as its occupants and neighborhood. To maximize the number of repeat observations, I use data from 1984 to 2011, a period of 27 years.

Following Rosen (1974), I control for all observable characteristics of the unit, household, and surrounding area which may affect the owner's self-reported value of their home or the annual contract rent that a tenant pays. This includes the number of rooms, number of bathrooms, the floor area of the unit, whether it is a detached or multiunit building, the presence of a working fireplace, and whether it includes a balcony or porch. Since certain features may be more desirable depending on the where the unit is located and furthermore, since there is large regional variation in my sample, I include several interaction effects to account for differences in the inherent value of amenities following Tsoodle and Turner (2008). These include: *air conditioning (ac)* × *hot*, *fireplace* × *cold*, and *parking* × *cold* where *cold* and *hot* are average temperature indicator variables following the AHS definition, *ac* is an indicator if the unit has central air conditioning, *parking* is an indicator for the presence of covered parking, and *fireplace* is an indicator for the existence of a working fireplace. It is plausible that households in areas with colder climates may value attributes such as a fireplace or covered parking more compared to those living in more temperate climates.

To account for variation in the heads of household, I utilize observable characteristics that I expect will influence the self-reported property value including whether the head of household is male, married, college-educated, and older than 65 years. Additionally, I control for household income and differences in observable housing quality. These include indicators for whether the unit has a missing roof, missing walls, broken windows, window bars, whether there have been leaks from the inside or outside in the past 30 days, whether the unit has cracks, and whether there is broken plaster. It should be noted that the wording of the AHS questions changed slightly after 1993 with a new survey format, although it is unlikely to have a significant impact on the results.[3] I also include self-reported measures of satisfaction with the unit and surrounding neighborhood which is included in the AHS. The question asks "On a scale from 1–10, how satisfied are you with the housing unit?" and similarly for the surrounding neighborhood.

In addition to the quality and satisfaction measures, I also include costs of annual routine maintenance and annual property taxes. Harding et al. (2000) find evidence that suggests homeowners may undermaintain housing if borrowers have limited liability in the case of mortgage default. Unfortunately, there are not equivalent survey questions for renters in the AHS data, and therefore I cannot control for these in the renter model specification. Besides maintenance and property taxes, I control for various geographic attributes, including the zone-level average household size to account for differences between suburban and urban areas.[4] I expect that larger households will tend to live in more suburban areas. I also include the average

[3]See the AHS codebook for further details.

[4]The AHS defines a zone as an economically homogenenous area with a population of 100,000 or more.

number of rooms, average household income, the fraction of black households, the fraction of households with a college education, and average building age, all at the zone-level, to capture neighborhood effects.

For the renter sample estimation I exclude all subsidized units (including government owned housing projects or rent-controlled units). These types of properties do not reflect the true market rent and hence, should not be impacted by the "shadow mortgage." I only consider occupied units in both the owner and renter samples. The justification for including both owner and renter-occupied units in this analysis is two-fold. Tiebout (1956) states that in regards to location decisions, people vote with their feet, although this effect will not likely be immediate. I include the renter-occupied specification since I expect that the decline in the willingness-to-pay for rental units should similarly be impacted by signs of fiscal stress.

In addition to the housing data from the AHS, I utilize data from the U.S. Census of Governments Annual Survey of State and Local Government Finances. This is an annual survey which assesses the financial state of governments across all levels, although not every government and government type is sampled in each survey.[5] I construct the solvency ratios using data from this survey. To obtain a wide and comprehensive assessment of municipal fiscal health, I employ several different measures. Following the literature, I use measures of cash, budgetary, service-level, and long-run solvency presented in Table 1.3 following Wang et al. (2007) and later utilized by Anders and Gearhart (2018). These ratios capture the municipal financial condition and can be useful in predicting fiscal stress. Gorina et al. (2018) find that the financial variables most associated with municipal bankruptcies are cash solvency, long-run solvency, and service-level solvency and can be significant in predicting bankruptcies. Alternative measures to predict the likelihood of a bankruptcy

[5] The full sample of local governments are only surveyed in years ending in a '2' or a '7'.

filing include a composite fiscal condition score, although the literature is mixed on this approach. McDonald (2017) deems it ineffective and Wang et al. (2007) finds that it is only marginally effective in predicting bankruptcies.

To capture the municipal fiscal state, I include several different measures presented in Table 1.3. Cash solvency is related to liquidity and demonstrates the ability of a local government to generate funds to pay its current liabilities. To assess the cash solvency, I utilize *Debt-to-Cash* (total debt outstanding to total cash and securities) which captures the ability of a government to balance its budget in the short-run. Smaller values of *Debt-to-Cash* indicate a municipality that is more cash solvent (all else equal). In addition to the cash solvency measure, I include the *Operating Ratio* (total revenue to total expenditures) which captures budgetary solvency. This ratio reflects the ability of a municipality to balance its budget. Increases in this ratio reflect a municipality that is better able to balance their budget, which I expect to have a positive and significant effect on housing values and furthermore, positively reflected in the rate of depreciation of housing. Although, this may be misleading since many local governments are subject to balanced-budget constraints.

I use four measures of service-level solvency including *EperC* (total expenditure to total population), *RperC* (total revenue to total population), *TperC* (total taxes to total population), and *Rev-to-CapOut* (total revenue to total capital outlay). The service-level solvency measures capture the ability of a municipality to provide basic services in the form of local amenities. I expect increases in *EperC* to positively impact housing values and result in lower rates of depreciation, all else equal. Larger values of this measure reflect a municipality with high levels of per capita spending on amenities such as parks and schools. To measure long-run solvency, I include *LTLiability* (total long-term liabilities to total interest on debt), *LTLiabilityperC* (total long-term liabilities to total population), and *Debt-to-Revenue* (total debt outstanding to total revenue). Both *LTLiability* and *LTLiabilityperC* capture current obligations as well as any other obligations that arise upon due date. I expect that increases in these measures will have a negative impact on housing values, as this suggests a government that is less financially stable.

Since houses are non-fungible and largely immobile, housing markets are tied to both the national and local level conditions. To control for local economic conditions, I obtain data from the U.S. Census and BLS. To control for any existing economic trends, I use unemployment data at the county-level. These data are from the BLS Local Area Unemployment Statistics. Note that since the BLS only began reporting unemployment rates at the county-level in 1990, I use the 1980 U.S. Census unemployment rate for observations appearing prior to 1985 and the 1990 value for observations appearing from 1985 to 1989. Furthermore, I control for both population and population change at the county-level using population estimates from the U.S. Census of Governments. This should account for any changes in household sorting over time.

Table 1.1 presents descriptive statistics by county, MSA, and tenure. Columns 1 and 2 list the county name and the MSA of each municipality in the sample, respectively. The year(s) that each county appear(s) in the sample is (are) listed in column 3. Counties are observed anywhere from one to five times, with 75 percent appearing three or more times in the sample. Columns 4–6 present the sample size by tenure group and list the maximum number of times a given household is observed. Over half of households appear three or more times in the sample. The sample size of owner-occupied properties versus renter-occupied properties varies by county and reflects the tenure breakdown for that location. Maricopa, Arizona has an owner sample of 5,641 households, the largest of any county. San Diego, California has the largest renter sample at 4,454 observations. Detailed descriptions of the housing variables and financial variables appear in Tables 1.2–1.3, respectively.

Selective descriptive statistics for the owner sample appear in Table 1.4. There is considerable variation in self-reported housing values with median and mean housing values of $149,986 and $210,406, reflecting a right-skewed distribution. The distribution of building ages is similarly skewed, with a median building age of 30 and a mean of over 34 years. The average number of rooms is seven and slightly over three percent of households reside in trailers. Around 87 percent of households reside in single-family detached houses, compared with just three percent that reside in multiunit structures, or buildings with more than four units. Notably, overall housing satisfaction is relatively high, with a median rating of nine, and an average of 8.5.

There is significant variation in the solvency measures. The first quartile *Debt-to-Cash* observation, a cash solvency measure, is 0.518. A ratio value of less than one may reflect a municipality with a large amount of cash reserves relative to outstanding short-term debt obligations. For the budgetary solvency measure *Operating Ratio*, the mean value is 1.032 which may reflect balanced budget constraints. The ratios with the the largest variation are *LTLiability*, a long-run solvency measure, and *Rev-to-CapOut*, a service-level solvency measure. Both capture long-term capital and infrastructure expenditures and are therefore subject to volatility since municipalities may not invest in these projects at the same rate across time. The first and third quartile of *LTLiability* are 12.931 and 20.315, respectively. For *Rev-to-CapOut*, the median and mean observation are 1,163.672 and 1,507.223, respectively, reflecting a right-skewed distribution. The solvency measure with the least variation is *TperC*, with a standard deviation of 0.540.

1.5 Empirical Results

The baseline model includes all controls discussed in the data section and presented in Table (1.2), excluding any solvency measures or interactions with solvency measures. The results from the baseline estimation for the owner sample are presented in Table 1.5. In column 1 of Table 1.5, I present results using the traditional approach, including both a linear and quadratic building age term, to capture the possible nonlinear effects of age on housing value. The coefficients on both the linear and quadratic building age terms are significant at one percent and have the expected signs, reflecting the nonlinear effect that building age has on the response variable. The results from the baseline estimation including the building age step function are presented in column 2. The coefficients suggest that the rate of depreciation is higher for relatively newer buildings and then tends to slow as the home ages, similar to findings from Yoshida (2016).

Table 1.6 shows the results controlling for *Debt-to-Cash*, a cash solvency measure. This measures captures outstanding debt obligations that are due in less than one year and larger values of this ratio suggest a municipality that is less liquid over the short-run. In column 1, results from the linear estimation suggest that a ten percent increase in *Debt-to-Cash* leads to a 0.07 percent decrease in housing values, roughly \$145. In column 2, a ten percent increase in *Debt-to-Cash* leads to a 0.17 percent decrease at the 0.1 quantile of housing values, over \$100. This is compared with a 0.02 increase, approximately \$70, at the 0.9 quantile in column 6. The results indicate that homeowners at the lower end of the distribution have a willingness-to-avoid increases in *Debt-to-Cash*. This seems to suggest that while these households bear the costs from from short-run capital investments such as highway and road improvements, they may not benefit from it. Doucet et al. (2011) discuss this new form of gentrification and suggest that certain areas may receive more targeted investment money and projects. This frequently occurs when a city is trying to attract and also keep more affluent residents.

I present results examining the impact of *Operating Ratio*, a budgetary solvency measure, in Table 1.7. Larger values of this measure reflect a local government that is more adept at balancing their budget and suggests greater fiscal health. While the main and interaction term coefficients are insignificant in the linear estimation, both are significant at one percent in the 0.1 and 0.25 quantile estimation in columns 2–3. In column 3, the marginal effect of *Operating Ratio* indicates that a ten percent increase in this measure leads to a 1.1 percent increase in housing values at the 0.25 quantile, over $1,100. The findings imply that increases in budgetary solvency have a positive impact on low-value properties, which reflects the government's ability to cover current or desired service levels. This could signal to households that a municipality may already be in a steady state of amenity provision. It should be noted that in columns 4–6, the marginal effect is also positive for high-value homeowners, but main coefficient term on *Operating Ratio* is not statistically significant at one percent.

To measure the service-level solvency of a municipality, several different measures are utilized. Table 1.8 presents the results controlling for *EperC*. Larger values of this measure reflect a local government that is better able to provide public amenities. The marginal effect suggests that a ten percent increase in *EperC* leads to over a 0.3 percent increase in the 0.25 quantile of housing values, or roughly $300. García et al. (2010) find a similar positive effect of increases in expenditures on housing values. In column 5, at the 0.75 quantile this leads to a 0.7 percent decrease in housing values, or around $1,800. At the 0.9 quantile, this results in roughly a 0.9 percent decrease, over $3,500. Greater service-level solvency may positively effect housing values at the lower end of the distribution due to the relative lack of privately provided amenities compared with those at the upper end. Nechyba and Walsh (2004) note that fiscal amenities matter and permit high-income households to escape redistributive taxation and improve public good quality.

The results from the estimation controlling for *TperC* are presented in Table 1.9, another service-level solvency measure. Larger values of *TperC*, suggest a more service-level solvent government. Oates (1969) finds that increases in the property tax rate tends to have a depressing effect on housing values, using the Tiebout sorting hypothesis to explain the finding. The marginal effect of *TperC* on the response variable is negative across all model estimations. In column 4, a ten percent increase in *TperC* leads to an approximate 0.4 percent decrease at the 0.5 quantile of housing values, approximately \$700. By contrast, at the 0.9 quantile in column 6, this results in a 0.9 percent decrease in housing values, around \$3,800. These results suggest that "house-rich" homeowners bear a larger burden from increases in this measure, but the effect is uniformly negative across the distribution of housing values.

In Table 1.10, I examine the impact of *LTLiability*, a long-run solvency measure. Larger values of measure suggest a municipality that is less solvent over the long-run. The marginal effect of *LTLiability* is negative across all model specifications, reflecting the shadow mortgage. At the 0.25 quantile of housing values, a ten percent increase in *LTLiability* leads to a 0.2 percent decrease, around \$200. In column 4, at the 0.5 quantile of housing values, a ten percent increase in *LTLiability* results in a 0.1 percent decrease, or over \$150. While these are relatively small effects, it should be noted that these financial measures are subject to volatility and it would not be unreasonable to see large changes over time. Hence, a ten percent increase is likely understating the true year-to-year change in the measure. Notably, the findings from MacKay (2014) show a 2.5–3.7 percent decrease in housing prices as a result of the negative news coverage regarding municipal debt obligations.

In Table 1.11, the implied variation in annual depreciation is estimated using the results from the linear estimation of (3.13) and evaluated at different percentiles of the solvency variables. Notably, the short and long-run solvency measures see the largest variation between the 10 and 90 percentiles. At the 90 percentile the implied depreciation for *Debt-to-Cash* is over 0.7 percent, compared with 0.3 percent at the 10 percentile, for a difference of nearly 0.4 percent annually. For *LTLiability*, the difference in implied annual depreciation between the 10 and 90 percentile is around 0.3 percent annually. Interestingly, higher-levels of service-level solvency tended to have a dampening effect on annual depreciation, with a difference of -0.20 and -0.28 for *TperC* and *EperC*, respectively. These findings indicate the higher levels of spending towards amenities may help slow the rate of depreciation of housing, whereas higher levels of debt, both over the short and long-run, tend to erode housing values more quickly.

Selected results from the estimation of the VC model for the owner sample are presented following the tables. They largely confirm the findings from the empirical results of the linear and quantile regressions. The estimation of Figure 1.1 suggests that the effect of *Debt-to-Cash* on housing values is positive for relatively new to average-aged buildings, but the relationship is nearly linearly negative. The effect of the *Operating Ratio* is positive across building ages, with wide confidence intervals, and has the strongest positive effect on the oldest structures in the sample. In Figure 1.3, the effect of *EperC* on housing values is highly nonlinear and largely negative, although there does appear to be a positive impact on housing values for the oldest properties in the sample. For *LTLiability* in Figure 1.5, the impact of this ratio on housing values is positive for buildings 20 years or younger. For older properties, increases in this measure negatively impact housing values.

Along with the main model specifications, I include additional owner sample model estimations as well as the full renter sample estimation in the separate appendix. The results for the renter sample estimation are very similar to the empirical findings of the owner sample estimation and suggest the theoretical model is appropriate to use for model predictions. Furthermore, the estimates of implied annual depreciation for the renter-occupied sample remained similar, with less variability, to the estimation using the owner sample. All the models estimated for the owner-occupied specification are similarly estimated for the renter sample, but with the noted differences as discussed in Section 1.3.

1.6 Discussion and Conclusion

This paper examines and provides evidence for capitalization of the shadow mortgage. For the average-aged building, increases in cash and long-run solvency measures, suggesting a less-solvent local government, negatively impact housing values. Specifically, a ten percent increase in total debt outstanding to total interest on debt results in an approximate 0.2 decline in housing values at the 0.1 quantile, roughly $100. By contrast, at the 0.9 quantile of housing values, a ten percent increase in this measure leads to a 0.02 percent, or around $70. While these effects are small for the individual homeowner, they can have significant impacts in aggregate. It should also be noted that due to the volatile nature of government investment, on average a ten percent change in these measures likely understates the true year-to-year change.

Unsurprisingly, increases in per capita expenditure positively impact housing values at the lower end of the distribution. A ten percent increase leads to a 0.6 percent increase, nearly $400, at the 0.1 quantile. Surprisingly, the results suggest that at the 0.9 quantile, this results in a 0.9 percent decrease, over $3,500. Notably, Skidmore and Scorsone (2011) find that in times of fiscal stress, municipalities respond by cutting back on recreation spending and capital improvements, as well as other maintenance projects. These findings suggest that local governments should focus on amenity-spending, even in economic downturns, to shield local housing values, especially low-value properties. In regards to taxation, Mieszkowski and Zodrow (1989) discuss two opposing views on the property tax. They consist of the benefit view, suggesting it is effectively a head tax, and the new view which suggests that capital bears the average burden of the tax. While I do not explicitly examine the impact of the property tax, my findings suggest that increases in taxes per capita has a uniformly negative impact on housing values, with high-value properties bearing a larger impact.

Moreover, Brueckner and Helsley (2011) illustrate how the market failures contributing to urban sprawl also impact and contribute to urban blight. They note that excessive suburban development can depress central-city housing and undermine maintenance incentives, leading to deficient levels of central-city investment. The policy implication would suggest shifting populations from the suburbs to city-center with reinvestments. This is further supported by the results from the estimation of depreciation which finds lower depreciation rates in areas with high per-capita expenditures, whereas municipalities with large cash and long-run measures have higher depreciation rates on average.

While gentrification is highly debated in local communities, Lang (1986) asserts that gentrification may promote municipal fiscal health if it increases the tax base by attracting affluent households. My findings are supported by previous studies including Hilber (2017) who find evidence that public and private investments and also intergovernmental transfers get capitalized into local house prices, especially in areas with strict regulatory and supply constraints. Local governments can benefit select groups by focusing their spending on certain areas as subject to their budgetary allowances. The policy implications of these findings may be useful to address urban issues through focused municipal spending.

Table 1.1: Housing Data Sample Statistics

County	MSA	Years	No. owners	No. renters	Max no. obs.
Adams, CO	Denver	86, 90, 95, 04, 11	1,116	454	3
Alameda, CA	San Francisco	85, 89, 93	1,185	1,134	3
Alexandria City, VA	Washington	85, 89, 93, 98, 07	125	224	3
Allegheny, PA	Pittsburgh	86, 90, 95, 04, 11	2,882	1,467	3
Arapahoe, CO	Denver	86, 90, 95, 04, 11	1,579	506	3
Arlington, VA	Washington	85, 89, 93, 98, 07	199	200	3
Ashtabula, OH	Cleveland	04	88	10	1
Baltimore City, MD	Baltimore	87, 91, 98, 07	644	496	2
Baltimore, MD	Baltimore	87, 91, 98, 07	1,197	506	2
Beaver, PA	Pittsburgh	86, 90, 95, 04, 11	478	119	3
Boulder, CO	Denver	86, 90	218	256	2
Brazoria, TX	Houston	87, 91	80	52	2
Broward, FL	Miami	86, 90, 95, 04, 11	1,980	1,075	3
Bucks, PA	Philadelphia	85, 89	298	156	2
Burlington, NJ	Philadelphia	85, 89	162	94	2
Butler, OH	Mansfield	11	195	26	1
Butler, PA	Pittsburgh	95, 04, 11	310	76	3
Camden, NJ	Philadelphia	85, 89	157	99	2
Chambers, TX	Houston	98	10	3	1
Chester, PA	Philadelphia	85, 89	127	95	2
Clark, WA	Seattle	86, 90, 95, 02	910	546	2
Clayton, GA	Atlanta	87, 91, 96, 04, 11	238	190	3
Cobb, GA	Atlanta	87, 91, 96, 04, 11	815	586	2
Collin, TX	Dallas	85, 89, 94, 02, 11	840	381	4
Contra Costa, CA	San Francisco	85, 89, 93	871	648	3
Cook, IL	Chicago	87, 91	931	698	2
Cuyahoga, OH	Cleveland	84, 88, 92, 96, 04, 11	4,608	1,280	3
Dade, FL	Miami	86	277	191	1
Dallas, TX	Dallas	85, 89, 94, 02, 11	3,146	3,382	4
Davis, UT	Salt Lake	84, 88, 92, 98	1,060	411	3
Delaware, PA	Philadelphia	85, 89	235	76	2
Denton, TX	Dallas	85, 89, 94, 02, 11	583	374	4
Denver, CO	Denver	86, 90, 95, 04, 11	1,518	1,318	3
Du Page, IL	Chicago	87, 91	140	111	2
Erie, NY	Buffalo	84, 88, 94, 02, 11	3,822	1,303	4
Fayette, PA	Pittsburgh	86, 90, 04, 11	236	83	2
Frederick, MD	Washington	85, 89, 93, 98, 07	237	77	3
Gloucester, NJ	Philadelphia	85, 89	75	32	2
Gwinnett, GA	Atlanta	87, 91, 96, 04, 11	811	393	3
Hamilton, OH	Cincinnati	86, 90, 98, 11	1,883	729	2
Harford, MD	Baltimore	87, 91, 98, 07	464	94	2
Hillsborough, FL	Tampa	85, 89, 93, 98, 07	1,903	1,396	3
Jackson, MO	Kansas City	86, 90, 95, 02	1,834	1,100	2
Jefferson, AL	Birmingham	84, 88, 92, 98, 11	3,745	2,373	3
Jefferson, MO	Kansas City	87, 91, 96, 04	388	94	2
Kane, IL	Chicago	87, 91	58	34	2
King, WA	Seattle	04, 09	1,700	768	2
Lake, IL	Chicago	87, 91	70	31	2
Livingston, MI	Detroit	85, 89, 93	69	29	2
Lorain, OH	Cleveland	11	160	36	1
Los Angeles, CA	Los Angeles	85, 89, 11	2,927	2,864	2
Macomb, MI	Detroit	85, 89, 93	551	182	3
Maricopa, AZ	Phoenix	85, 89, 94, 02, 11	5,641	4,276	4
Marin, CA	San Francisco	85, 89, 93, 98, 11	600	458	3
McHenry, IL	Chicago	91	25	16	1
Milwaukee, WI	Milwaukee	84, 88, 94, 02, 11	2,656	2,202	4
Monroe, MI	Detroit	85, 89, 93	89	33	3
Monroe, NY	Buffalo	86, 90, 98	2,185	955	2
Montgomery, MD	Washington	85, 89, 93, 98, 07	1,027	395	3
Montgomery, PA	Philadelphia	85, 89	323	166	2
Newport News City, VA	Norfolk	84, 88, 92, 98	550	387	3
Niagara, NY	Buffalo	84, 88, 94, 02, 11	912	416	4
Norfolk City, VA	Norfolk	84, 88, 92	522	522	3

Housing Data Sample Statistics — continued

County	MSA	Years	No. owners	No. renters	Max no. obs.
Oakland, MI	Detroit	85, 89, 93	637	562	3
Orange, CA	Anaheim	86, 90, 94, 02, 11	5,418	4,418	4
Palm Beach, FL	Miami	07	173	32	1
Philadelphia, PA	Philadelphia	85, 89	269	221	2
Pierce, WA	Seattle	87, 91, 09	615	525	2
Pinal, AZ	Phoenix	11	72	17	1
Pinellas, FL	Tampa	85, 89, 93, 98, 07	2,142	1,505	3
Portsmouth City, VA	Norfolk	84, 88, 92	282	177	3
Prince Georges, MD	Washington	85, 89, 93, 98, 07	931	411	3
Riverside, CA	San Bernardino	86, 90, 94, 02	2,224	1,618	4
Salt Lake, UT	Salt Lake	84, 88, 92, 98	4,086	2,645	3
San Bernardino, CA	San Bernardino	86, 90, 94, 02	2,228	2,110	4
San Diego, CA	San Diego	91, 94, 02, 11	4,434	4,454	3
San Francisco, CA	San Francisco	85, 89, 93, 98, 11	811	1,722	3
San Mateo, CA	San Francisco	85, 89, 93, 98, 11	1,710	975	3
Snohomish, WA	Olympia	09	55	12	1
St. Louis City, MO	St. Louis	87, 91, 96, 04, 11	453	372	3
St. Louis, MO	St. Louis	87, 91, 96, 04, 11	2,523	801	3
Virginia Beach City, VA	Norfolk	84, 88, 92, 98	1,559	928	3
Washington, DC	Washington	85, 89, 93, 98, 07	302	398	3
Washington, PA	Pittsburgh	86, 90, 95, 04	372	90	2
Waukesha, WI	Milwaukee	84, 88, 94, 02, 11	1,554	375	4
Wayne, MI	Detroit	85, 89, 93	1,309	781	3
Weber, UT	Salt Lake	88, 92, 98	732	353	2
Westmoreland, PA	Pittsburgh	86, 90, 95, 04, 11	1,030	208	3

Table 1.2: Description of Variables

Variable Name	Variable Description
Dependent Variable	
propvalue	self-reported current market value ($)
annualrent	annual contract rent ($)
Unit Characteristics	
sqfootage	area of unit (1000s of sq ft)
proptax	annual property taxes (owner-sample only) ($)
annualmain	cost of annual maintenance (owner-sample only) ($)
numrooms	number of rooms in unit
age	age of unit at time of observation
numbaths	number of full bathrooms
single	indicator if detached unit
mobile	indicator if mobile home/trailer
large	indicator if located in building with more than four units
centralac	indicator if central ac present
centralheat	indicator if central heat present
fireplace	indicator if working fireplace
balcony	indicator if unit has a porch/balcony
hquality	self-reported housing satisfaction rating
nquality	self-reported nbrhood satisfaction rating
parking	indicator if unit has covered parking
hot	indicator if unit located in county described as “hot” climate (AHS 2011 description)
cold	indicator if unit located in county described as “coldest” climate (AHS 2011 description)
Household Characteristics	
male	indicator if male HOH
married	indicator if married HOH
black	indicator if black HOH
college	indicator if four years of college or more completed by HOH
older	indicator if HOH is 65 years or older
hhincome	total household income ($)
Local-Level Controls	
avgnumper	zone-level average of household size
avgnumrooms	zone-level average number of rooms
avghhincome	zone-level average household income ($)
avgblack	zone-level fraction of black HOHs
avgcollege	zone-level fraction with four years of college or more
avgage	zone-level average age of building
unemprate	county-level unemployment rate
popchg	county-level population change
population	county-level population
Geographic Controls	
state	state (FIPS code)
centralcity	indicator if unit is in central city of MSA
MSA	metropolitan statistical area of unit location
county	county (FIPS code)
zone	socio-economically homogeneous area with 100K population or more

Table 1.3: Description of Financial Ratios

Solvency Measure	Description	Financial Ratio	Solvency Type
Debt-to-Cash	Total debt outstanding to total cash and securities	$\frac{\text{Total debt outstanding}}{\text{Total cash and securities}}$	Cash
Operating Ratio	Total revenue to total expenditure	$\frac{\text{Total revenue}}{\text{Total expenditure}}$	Budgetary
Debt-to-Revenue	Total debt outstanding to total revenue	$\frac{\text{Total debt outstanding}}{\text{Total revenue}}$	Long-run
LTLiability	Total debt due within more than one year to total interest on debt	$\frac{\text{Total long-term debt outstanding}}{\text{Total interest on debt}}$	Long-run
LTLiabilityperC	Total debt due within more than one year to total population	$\frac{\text{Total long-term debt outstanding}}{\text{Total population}}$	Long-run
EperC	Total expenditure to total population	$\frac{\text{Total expenditures}}{\text{Total population}}$	Service-level
TperC	Total taxes to total population	$\frac{\text{Total taxes}}{\text{Total population}}$	Service-level
RperC	Total revenue to total population	$\frac{\text{Total revenue}}{\text{Total population}}$	Service-level
Rev-to-CapOut	Total revenue to total expenditure on capital	$\frac{\text{Total revenue}}{\text{Total capital outlay}}$	Service-level

*All financial variables are for prior fiscal year.

Table 1.4: Descriptive Statistics for Owner Sample

	— Owner Sample —				
Variable	Q1	Median	Mean	Q3	Std. Dev.
Unit & Household					
Building age	17	30	33.712	48	21.294
Number of rooms	5	6	6.576	8	1.717
Number of bathrooms	1	2	1.696	2	0.718
Trailer	0	0	0.032	0	0.175
Fireplace	0	1	0.530	1	0.499
Balcony	1	1	0.897	1	0.303
Central ac	0	1	0.556	1	0.497
Central heat	1	1	0.751	1	0.432
Detached structure	1	1	0.869	1	0.338
Large structure	0	0	0.026	0	0.158
Housing satisfaction	8	9	8.484	10	1.491
Neighborhood satisfaction	7	8	8.140	10	1.796
Missing roof	0	0	0.021	0	0.143
Visible cracks	0	0	0.041	0	0.198
Broken windows	0	0	0.018	0	0.133
Covered parking	1	1	0.840	1	0.367
Housing value ($)	98533.160	149986.400	210406.300	247569.400	215884.900
Log housing value ($)	11.498	11.918	12.257	12.419	0.926
Log square footage	0.262	0.588	0.577	0.875	0.501
Log household income ($)	10.593	11.115	11.024	11.542	0.876
Log population	13.387	13.745	13.782	14.360	0.871
Financial Ratios					
Debt-to-Cash	0.518	0.859	1.089	1.282	1.198
STDebt-to-Cash	0	0	0.051	0.011	0.185
Operating Ratio	0.982	1.031	1.032	1.078	0.106
LT Liability	12.931	15.877	18.663	20.315	24.975
LTLiabilityperC	0.396	0.722	1.275	1.531	1.778
Debt-to-Revenue	0.470	0.797	0.968	1.212	0.744
TperC	0.222	0.319	0.482	0.510	0.540
RperC	0.655	1.102	1.374	1.618	1.321
EperC	0.637	1.084	1.332	1.566	1.228
Rev-to-CapOut	706.854	1163.672	1507.223	1739.885	1458.572

*All dollar values are in 2010 U.S. dollars.

Table 1.5: Baseline Linear Model Estimates

Dependent Variable: Log housing value	— **Owner Sample** — (1) Linear	(2) Linear
Building Age	-0.0082***	
	(0.0012)	
Building Age^2	5.32×10^{-05}***	
	(1.20×10^{-05})	
Building Age		
$\times$ $\mathbb{I}$(0 - 8 years)		-0.0131***
		(0.0028)
$\times$ $\mathbb{I}$(9 - 17 years)		-0.0073***
		(0.0013)
$\times$ $\mathbb{I}$(18 - 26 years)		-0.0061***
		(0.0009)
$\times$ $\mathbb{I}$(27 - 35 years)		-0.0047***
		(0.0007)
$\times$ $\mathbb{I}$(36 - 44 years)		-0.0036***
		(0.0006)
$\times$ $\mathbb{I}$(45 - 53 years)		-0.0029***
		(0.0005)
$\times$ $\mathbb{I}$(54 - 62 years)		-0.0023***
		(0.0005)
$\times$ $\mathbb{I}$(63 - 71 years)		-0.0024***
		(0.0005)
$\times$ $\mathbb{I}$(72 - 80 years)		-0.0015***
		(0.0005)
$\times$ $\mathbb{I}$(80 years +)		-0.0008
		(0.0006)

This table presents selected results from the baseline estimation of (3.12) without controlling for any financial measures. The sample consists of 100,577 observations from 1984 to 2011 and includes 88 counties in total. All models include the full set of controls, year and location fixed effects, and building age is interacted with an indicator if the unit is detached and the floor area of the unit (logged and demeaned). In columns 1–2, the adjusted R-squared values are 0.6021 and 0.6022, respectively. Standard errors are in parentheses and are clustered at the county-level using AHS sample weights. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1.6: Linear and Quantile Model Estimates – Cash Solvency

Dependent Variable:	— **Owner Sample** —					
Log housing value	(1) Linear	(2) 0.1 quantile	(3) 0.25 quantile	(4) 0.5 quantile	(5) 0.75 quantile	(6) 0.9 quantile
Log *Debt-to-Cash*	0.0537**	0.0650***	0.0667***	0.0676***	0.0594***	0.0565***
	(0.0206)	(0.0089)	(0.0052)	(0.0051)	(0.0063)	(0.0086)
Building Age						
× Log *Debt-to-Cash*	−0.0018***	−0.0024***	−0.0022***	−0.0020***	−0.0017***	−0.0016***
	(0.0005)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Building Age						
× $\mathbb{I}$(0 - 8 years)	−0.0123***	−0.0109***	−0.0099***	−0.0098***	−0.0119***	−0.0118***
	(0.0028)	(0.0023)	(0.0013)	(0.0015)	(0.0016)	(0.0025)
× $\mathbb{I}$(9 - 17 years)	−0.0070***	−0.0099***	−0.0083***	−0.0075***	−0.0079***	−0.0090***
	(0.0012)	(0.0011)	(0.0006)	(0.0006)	(0.0008)	(0.0010)
× $\mathbb{I}$(18 - 26 years)	−0.0059***	−0.0093***	−0.0075***	−0.0064***	−0.0061***	−0.0055***
	(0.0009)	(0.0008)	(0.0004)	(0.0004)	(0.0005)	(0.007)
× $\mathbb{I}$(27 - 35 years)	−0.0046***	−0.0085***	−0.0066***	−0.0053***	−0.0046***	−0.0041***
	(0.0007)	(0.0007)	(0.0004)	(0.0004)	(0.0004)	(0.0006)
× $\mathbb{I}$(36 - 44 years)	−0.0035***	−0.0084***	−0.0062***	−0.0046***	−0.0037***	−0.0028***
	(0.0006)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
× $\mathbb{I}$(45 - 53 years)	−0.0030***	−0.0083***	−0.0061***	−0.0042***	−0.0031***	−0.0020***
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(54 - 62 years)	−0.0025***	−0.0084***	−0.0058***	−0.0038***	−0.0024***	−0.0012**
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(63 - 71 years)	−00025***	−0.0085***	−0.0059***	−0.0036***	−0.0020***	−0.0008*
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(72 - 80 years)	−0.0016***	−0.0084***	−0.00054***	−0.0030***	−0.0013***	0.0001
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(80 years +)	−0.0010	−0.0082***	−0.0052***	−0.0025***	−0.0004	0.0008
	(0.0007)	(0.0006)	(0.0004)	(0.0003)	(0.0001)	(0.0002)
Marginal Effect: Log *Debt-to-Cash*						
At Mean(Building Age):	−0.0069	−0.0176	−0.0078	−0.0004	0.0008	0.0017
	(0.0115)	(0.0078)	(0.0042)	(0.0040)	(0.0046)	(0.0066)

This table presents selected results from the estimation of (3.12) controlling for cash solvency. The cash solvency measure is Debt-to-Cash (total debt outstanding to total cash and securities). The sample consists of 100,557 observations from 1984 to 2011 and includes 88 counties in total. All models include the full set of controls, year and location fixed effects, and building age is interacted with an indicator if the unit is detached, the floor area of the unit (logged and demeaned), and Debt-to-Cash (logged and demeaned). The marginal effect of log Debt-to-Cash on log property value is estimated at the mean building age. In column 1, the adjusted R-squared value is 0.6033. For the linear specification, standard errors are in parentheses and are clustered at the county-level using AHS sample weights. For the quantile regression estimates, bootstrapped standard errors are in parentheses and are clustered at the county-level. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1.7: Linear and Quantile Model Estimates – Budgetary Solvency

Dependent Variable: Log housing value	(1) Linear	(2) 0.1 quantile	(3) 0.25 quantile	(4) 0.5 quantile	(5) 0.75 quantile	(6) 0.9 quantile
	— Owner Sample —					
Log *Operating Ratio*	0.0245	−0.2204***	−0.1454***	−0.0875***	0.0230	−0.0505
	(0.1359)	(0.0553)	(0.0320)	(0.0285)	(0.0341)	(0.0549)
Building Age						
× Log *Operating Ratio*	0.0040	0.0092***	0.0076***	0.0053***	0.0025***	0.0033**
	(0.0040)	(0.0015)	(0.0008)	(0.0008)	(0.0009)	(0.0014)
Building Age						
× $\mathbb{I}$(0 - 8 years)	−0.0130***	−0.0114***	−0.0111***	−0.0107***	−0.0125***	−0.0123***
	(0.0029)	(0.0023)	(0.0015)	(0.0015)	(0.0016)	(0.0023)
× $\mathbb{I}$(9 - 17 years)	−0.0072***	−0.0100***	−0.0089***	−0.0081***	−0.0085***	−0.0090***
	(0.0013)	(0.0010)	(0.0007)	(0.0007)	(0.0007)	(0.0010)
× $\mathbb{I}$(18 - 26 years)	−0.0060***	−0.0093***	−0.0078***	−0.0069***	−0.0064***	−0.0057***
	(0.0090)	(0.0007)	(0.0005)	(0.0005)	(0.0005)	(0.0007)
× $\mathbb{I}$(27 - 35 years)	−0.0046***	−0.0085***	−0.0069***	−0.0058***	−0.0049***	−0.0043***
	(0.0072)	(0.0006)	(0.0004)	(0.0004)	(0.0004)	(0.0006)
× $\mathbb{I}$(36 - 44 years)	−0.0035***	−0.0083***	−0.0064***	−0.0049***	−0.0038***	−0.0029***
	(0.0006)	(0.0006)	(0.0004)	(0.0004)	(0.0004)	(0.0006)
× $\mathbb{I}$(45 - 53 years)	−0.0028***	−0.0082***	−0.0062***	−0.0045***	−0.0032***	−0.0020***
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0006)
× $\mathbb{I}$(54 - 62 years)	−0.0023***	−0.0083***	−0.0059***	−0.0040***	−0.0024***	−0.0013**
	(0.0005)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(63 - 71 years)	−0.0024***	−0.0084***	−0.0061***	−0.0040***	−0.0021***	−0.0010*
	(0.0005)	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(72 - 80 years)	−0.0015***	−0.0086***	−0.0055***	−0.0032***	−0.0013***	0.0001
	(0.0005)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(80 years +)	−0.0007	−0.0078***	−0.0050***	−0.0026***	−0.0004	0.0008
	(0.0006)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
Marginal Effect: Log *Operating Ratio*						
At Mean(Building Age):	0.1595	0.0929	0.1144	0.0933	0.1070	0.0632
	(0.0929)	(0.0437)	(0.0227)	(0.0176)	(0.0216)	(0.0418)

This table presents selected results from the estimation of (3.12) controlling for budgetary solvency. The budgetary solvency measure is the Operating Ratio (total revenue to total expenditure). The sample consists of 100,557 observations from 1984 to 2011 and includes 88 counties in total. All models include the full set of controls, year and location fixed effects, and building age is interacted with an indicator if the unit is detached, the floor area of the unit (logged and demeaned), and the Operating Ratio (logged and demeaned). The marginal effect of log Operating Ratio on log property value is estimated at the mean building age. In column 1, the adjusted R-squared value is 0.6024. For the linear specification, standard errors are in parentheses and are clustered at the county-level using AHS sample weights. For the quantile regression estimates, bootstrapped standard errors are in parentheses and are clustered at the county-level. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1.8: Linear and Quantile Model Estimates – Service-level Solvency I

Dependent Variable:	— **Owner Sample** —					
Log housing value	(1) Linear	(2) 0.1 quantile	(3) 0.25 quantile	(4) 0.5 quantile	(5) 0.75 quantile	(6) 0.9 quantile
Log *EperC*	−0.0126***	−0.0258	−0.0453**	−0.0827***	−0.1171***	−0.1109***
	(0.0029)	(0.0296)	(0.0186)	(0.0151)	(0.0180)	(0.0259)
Building Age						
× Log *EperC*	−0.0070***	0.0025***	0.0022***	0.0019***	0.0013***	0.0007***
	(0.0012)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Building Age						
× $\mathbb{I}$(0 - 8 years)	−0.0126***	−0.0123***	−0.0113***	−0.0111***	−0.0126***	−0.0144***
	(0.0029)	(0.0021)	(0.0014)	(0.0015)	(0.0018)	(0.0024)
× $\mathbb{I}$(9 - 17 years)	−0.0070***	−0.0104***	−0.0092***	−0.0084***	−0.0085***	−0.0097***
	(0.0012)	(0.0010)	(0.0007)	(0.0007)	(0.0007)	(0.0010)
× $\mathbb{I}$(18 - 26 years)	−0.0058***	−0.0095***	−0.0082***	−0.0072***	−0.0065***	−0.0063***
	(0.0008)	(0.0007)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
× $\mathbb{I}$(27 - 35 years)	−0.0043***	−0.0085***	−0.0072***	−0.0059***	−0.0050***	−0.0047***
	(0.0007)	(0.0006)	(0.0004)	(0.0004)	(0.0004)	(0.0006)
× $\mathbb{I}$(36 - 44 years)	−0.0032***	−0.0083***	−0.0066***	−0.0050***	−0.0040***	−0.0034***
	(0.0006)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
× $\mathbb{I}$(45 - 53 years)	−0.0025***	−0.0083***	−0.0065***	−0.0046***	−0.0033***	−0.0025***
	(0.0005)	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0006)
× $\mathbb{I}$(54 - 62 years)	−0.0021***	−0.0083***	−0.0063***	−0.0042***	−0.0026***	−0.0017***
	(0.0005)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
× $\mathbb{I}$(63 - 71 years)	−0.0021***	−0.0086***	−0.0065***	−0.0042***	−0.0023***	−0.0013**
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(72 - 80 years)	−0.0015***	−0.0089***	−0.0063***	−0.0038***	−0.0017***	−0.0004
	(0.0005)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
× $\mathbb{I}$(80 years +)	−0.0009	−0.0086***	−0.0060***	−0.0032***	−0.0010***	0.0002***
	(0.0006)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
Marginal Effect: Log *EperC*						
At Mean(Building Age):	−0.0592	0.0609	0.0313	0.0196	−0.0714	−0.0859
	(0.0665)	(0.0288)	(0.0179)	(0.0149)	(0.0174)	(0.0242)

This table presents selected results from the estimation of (3.12) controlling for service-level solvency. The service-level solvency measure is EperC (total expenditure to population). The sample consists of 100,577 observations from 1984 to 2011 and includes 88 counties in total. All models include the full set of controls, year and location fixed effects, and building age is interacted with an indicator if the unit is detached, the floor area of the unit (logged and demeaned), and the EperC (logged and demeaned). The marginal effect of log EperC on log property value is estimated at the mean building age. In column 1, the adjusted R-squared value is 0.6028. For the linear specification, standard errors are in parentheses and are clustered at the county-level using AHS sample weights. For the quantile regression estimates, bootstrapped standard errors are in parentheses and are clustered at the county-level. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1.9: Linear and Quantile Model Estimates – Service-level Solvency II

Dependent Variable:	— **Owner Sample** —					
Log housing value	(1) Linear	(2) 0.1 quantile	(3) 0.25 quantile	(4) 0.5 quantile	(5) 0.75 quantile	(6) 0.9 quantile
Log *TperC*	−0.0790	−0.0739***	−0.0923***	−0.0904***	−0.0969***	−0.1046***
	(0.0520)	(0.0215)	(0.0125)	(0.0098)	(0.0144)	(0.0231)
Building Age						
× Log *TperC*	0.0012**	0.0019***	0.0016***	0.0014***	0.0010***	0.0004*
	(0.0005)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Building Age						
× $\mathbb{I}$(0 - 8 years)	−0.0127***	−0.0112***	−0.0111***	−0.0110***	−0.0122***	−0.0137***
	(0.0029)	(0.0022)	(0.0014)	(0.0015)	(0.0016)	(0.0023)
× $\mathbb{I}$(9 - 17 years)	−0.0072***	−0.0099***	−0.0093***	−0.0086***	−0.0087***	−0.0096***
	(0.0012)	(0.0010)	(0.0007)	(0.0006)	(0.0007)	(0.0010)
× $\mathbb{I}$(18 - 26 years)	−0.0059***	−0.0092***	−0.0083***	−0.0074***	−0.0066***	−0.0062***
	(0.0008)	(0.0007)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
× $\mathbb{I}$(27 - 35 years)	−0.0045***	−0.0083***	−0.0073***	−0.0060***	−0.0051***	−0.0046***
	(0.0007)	(0.0006)	(0.0004)	(0.0004)	(0.0004)	(0.0006)
× $\mathbb{I}$(36 - 44 years)	−0.0033***	−0.0082***	−0.0068***	−0.0052***	−0.0040***	−0.0032***
	(0.0006)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
× $\mathbb{I}$(45 - 53 years)	−0.0026***	−0.0081***	−0.0066***	−0.0047***	−0.0034***	−0.0023***
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0006)
× $\mathbb{I}$(54 - 62 years)	−0.0022***	−0.0082***	−0.0063***	−0.0043***	−0.0027***	−0.0015***
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(63 - 71 years)	−0.0022***	−0.0085***	−0.0065***	−0.0043***	−0.0024***	−0.0011**
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(72 - 80 years)	−0.0016***	−0.0088***	−0.0063***	−0.0039***	−0.0018***	−0.0003
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(80 years +)	−0.0008	−0.0082***	−0.0058***	−0.0032***	−0.0009**	0.0005
	(0.0006)	(0.0007)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
Marginal Effect:						
Log *TperC*						
At Mean(Building Age):	−0.0377	−0.0102	−0.0373	−0.0440	−0.0642	−0.0920
	(0.0478)	(0.0201)	(0.0115)	(0.0093)	(0.0143)	(0.0233)

This table presents selected results from the estimation of (3.12) controlling for service-level solvency. The service-level solvency measure is TperC (total taxes to population). The sample consists of 100,557 observations from 1984 to 2011 and includes 88 counties in total. All models include the full set of controls, year and location fixed effects, and building age is interacted with an indicator if the unit is detached, the floor area of the unit (logged and demeaned), and TperC (logged and demeaned). The marginal effect of log TperC on log property value is estimated at the mean building age. In column 1, the adjusted R-squared value is 0.6026. For the linear specification, standard errors are in parentheses and are clustered at the county-level using AHS sample weights. For the quantile regression estimates, bootstrapped standard errors are in parentheses and are clustered at the county-level. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1.10: Linear and Quantile Model Estimates – Long-run Solvency

Dependent Variable:	— **Owner Sample** —					
Log housing value	(1) Linear	(2) 0.1 quantile	(3) 0.25 quantile	(4) 0.5 quantile	(5) 0.75 quantile	(6) 0.9 quantile
Log *LTLiability*	0.0847**	0.0886***	0.0876***	0.1063***	0.1079***	0.1036***
	(0.0009)	(0.0151)	(0.0102)	(0.0089)	(0.0106)	(0.0158)
Building Age						
× Log *LTLiability*	−0.0029***	−0.0031***	−0.0032***	−0.0034***	−0.0032***	−0.0029***
	(0.0009)	(0.0004)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Building Age						
× $\mathbb{I}$(0 - 8 years)	−0.0132***	−0.0115***	−0.0126***	−0.0110***	−0.0129***	−0.0132***
	(0.0028)	(0.0024)	(0.0015)	(0.0014)	(0.0015)	(0.0022)
× $\mathbb{I}$(9 - 17 years)	−0.0074***	−0.0102***	−0.0095***	−0.0081***	−0.0086***	−0.0096***
	(0.0012)	(0.0011)	(0.0007)	(0.0006)	(0.0007)	(0.0010)
× $\mathbb{I}$(18 - 26 years)	−0.0061***	−0.0093***	−0.0083***	−0.0069***	−0.0065***	−0.0060***
	(0.0009)	(0.0008)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
× $\mathbb{I}$(27 - 35 years)	−0.0047***	−0.0085***	−0.0072***	−0.0056***	−0.0050***	−0.0044***
	(0.0007)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
× $\mathbb{I}$(36 - 44 years)	−0.0036***	−0.0083***	−0.0066***	−0.0048***	−0.0039***	−0.0031***
	(0.0006)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0006)
× $\mathbb{I}$(45 - 53 years)	−0.0028***	−0.0081***	−0.0063***	−0.0041***	−0.0032***	−0.0021***
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0006)
× $\mathbb{I}$(54 - 62 years)	−0.0023***	−0.0081***	−0.0059***	−0.0036***	−0.0025***	−0.0012**
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0006)
× $\mathbb{I}$(63 - 71 years)	−0.0024***	−0.0083***	−0.0061***	−0.0037***	−0.0023***	−0.0009*
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(72 - 80 years)	−0.0015***	−0.0085***	−0.0055***	−0.0030***	−0.0013***	0.0002
	(0.0005)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0005)
× $\mathbb{I}$(80 years +)	−0.0004	−0.0074***	−0.0047***	−0.0018***	0.0000	0.0011**
	(0.0006)	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
Marginal Effect:						
Log *LTLiability*						
At Mean(Building Age):	−0.0146	−0.0177	−0.0222	−0.0102	−0.0022	0.0033
	(0.0261)	(0.0103)	(0.0033)	(0.0056)	(0.0067)	(0.0193)

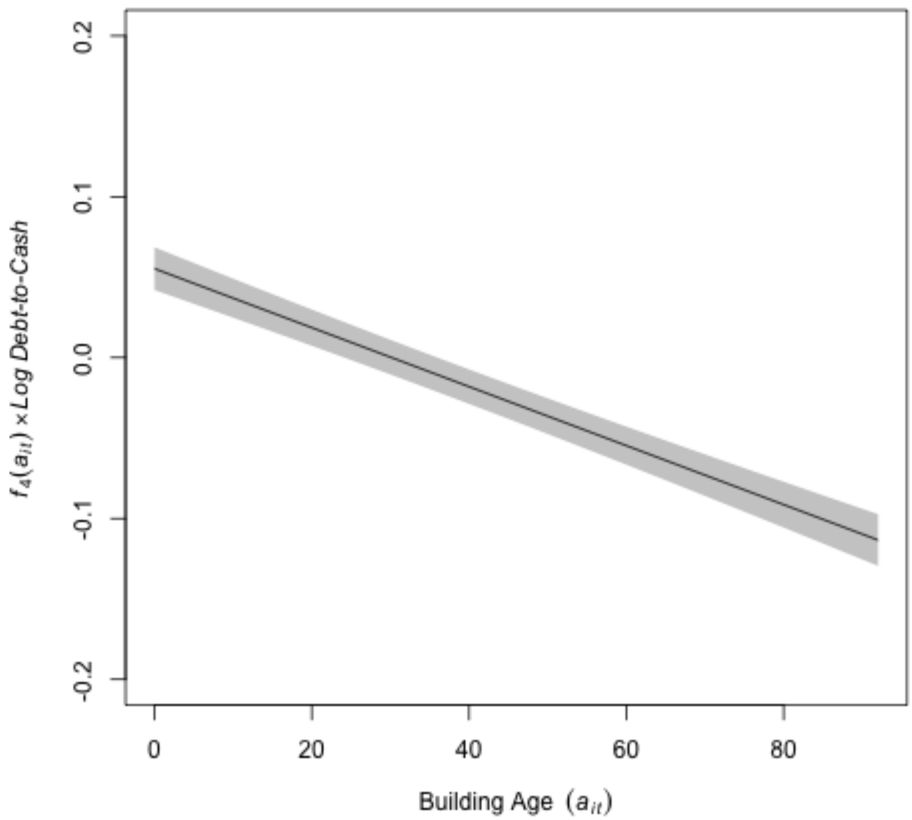
This table presents selected results from the estimation of (3.12) controlling for long-run solvency. The long-run solvency measure is LTLiability (total long-term debt outstanding to total interest on debt). The sample consists of 100,557 observations from 1984 to 2011 and includes 88 counties in total. All models include the full set of controls, year and location fixed effects, and building age is interacted with an indicator if the unit is detached, the floor area of the unit (logged and demeaned), and the LTLiability (logged and demeaned). The marginal effect of log LTLiability on log property value is estimated at the mean building age. In column 1, the adjusted R-squared value is 0.6029. For the linear specification, standard errors are in parentheses and are clustered at the county-level using AHS sample weights. For the quantile regression estimates, bootstrapped standard errors are in parentheses and are clustered at the county-level. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1.11: Variation in Implied Annual Depreciation

	— **Owner Sample** —					
	(1) 10 percentile	(2) 25 percentile	(3) 50 percentile	(4) 75 percentile	(5) 90 percentile	(6) 90 − 10 percentile
Log *Debt-to-Cash*	0.3332	0.4974	0.5865	0.6557	0.7291	0.3959
	(0.0011)	(0.0009)	(0.0008)	(0.0008)	(0.0009)	
Log *Operating Ratio*	0.6497	0.6162	0.5939	0.5738	0.5413	−0.1084
	(0.0009)	(0.0008)	(0.0008)	(0.0009)	(0.0011)	
Log *EperC*	0.7410	0.6851	0.5826	0.5235	0.4635	−0.2775
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0009)	
Log *TperC*	0.6734	0.6577	0.6099	0.5579	0.4713	−0.2021
	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0009)	
Log *LTLiability*	0.4634	0.5262	0.5992	0.6724	0.7355	0.2721
	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0009)	

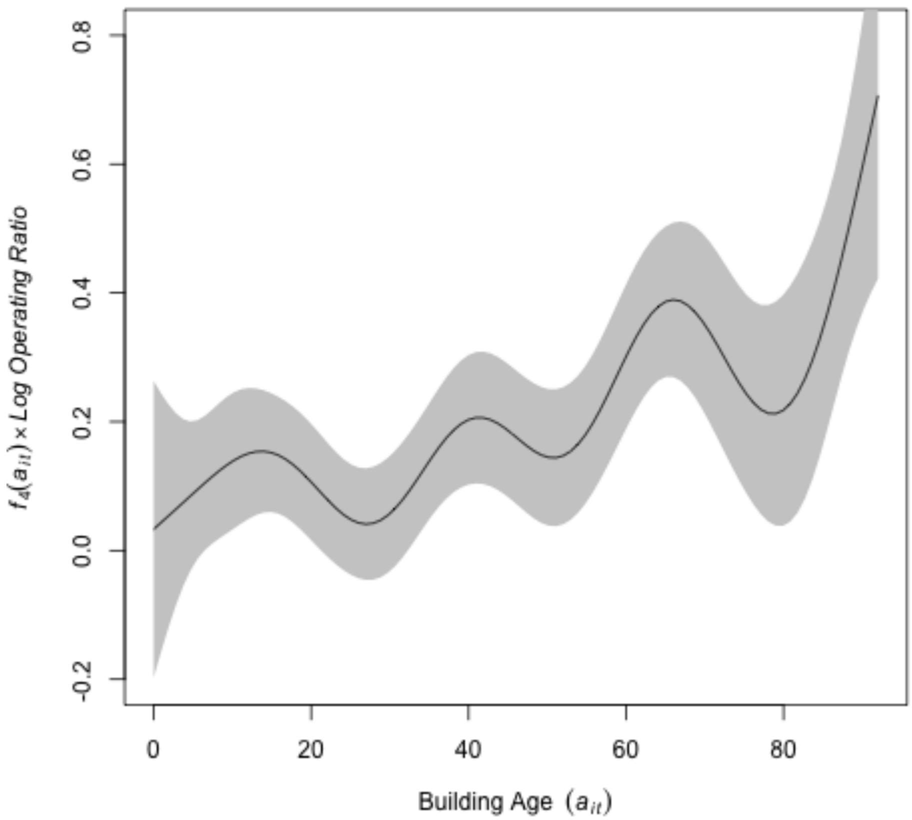
This table presents results from the estimation of (3.13) for each linear model in Tables 1.6–1.10, respectively, examining the marginal effect of building age at different percentiles of the solvency variable of interest. The presented values are annual depreciation rates in percentages (with negative values being appreciation). The difference in implied annual depreciation between the 90 and 10 percentile of each solvency variable is presented in column 6. Robust standard errors in parentheses.

Figure 1.1: Nonparametric Building Age Interaction with Debt-to-Cash



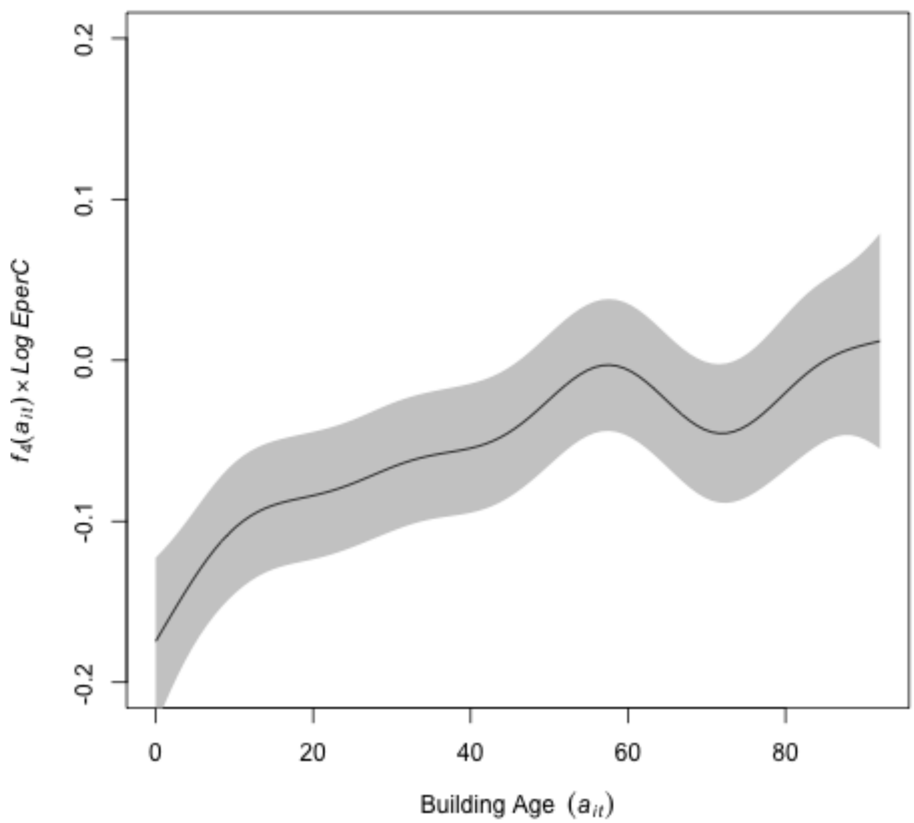
This figure presents results from the estimation of (3.17) controlling for Debt-to-Cash (total debt outstanding to total cash and securities) with 95% point-wise confidence interval. Building age is estimated nonparametrically using a thin plate spline and interacted with Debt-to-Cash (logged and demeaned), an indicator if the unit is detached, and the floor area (logged and demeaned). All smooth terms are significant at 1% with p-values approximately less than 0.001.

Figure 1.2: Nonparametric Building Age Interaction with the Operating Ratio



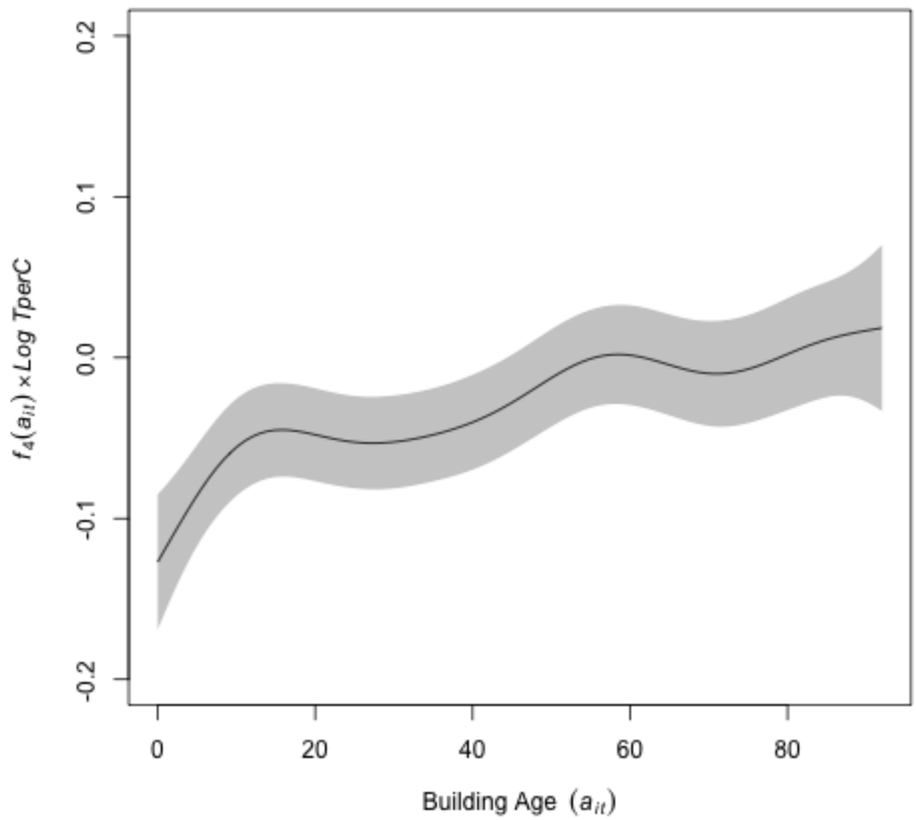
This figure presents results from the estimation of (3.17) controlling for Operating Ratio (total revenue to total expenditure) with 95% point-wise confidence interval. Building age is estimated nonparametrically using a thin plate spline and interacted with the Operating Ratio (logged and demeaned), an indicator if the unit is detached, and the floor area (logged and demeaned). All smooth terms are significant at 1% with p-values approximately less than 0.001. Note the scale change from other figures.

Figure 1.3: Nonparametric Building Age Interaction with EperC



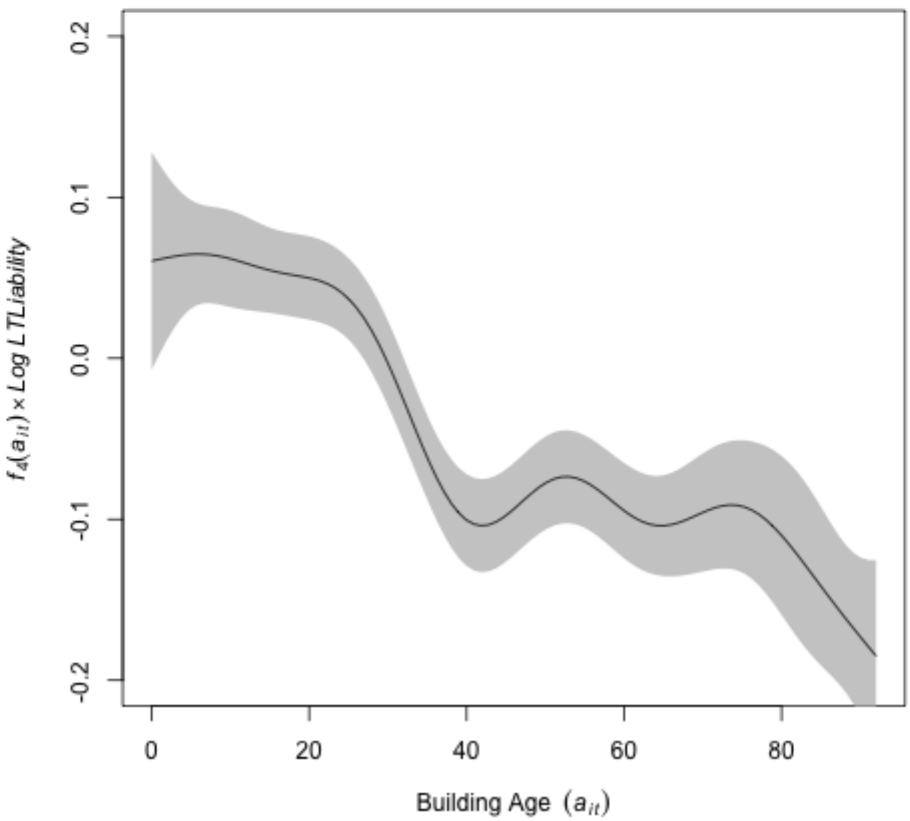
This figure presents results from the estimation of (3.17) controlling for EperC (total expenditure to total population) with 95% point-wise confidence interval. Building age is estimated nonparametrically using a thin plate spline and interacted with the EperC (logged and demeaned), an indicator if the unit is detached, and the floor area (logged and demeaned). All smooth terms are significant at 1% with p-values approximately less than 0.001.

Figure 1.4: Nonparametric Building Age Interaction with TperC



This figure presents results from the estimation of (3.17) controlling for TperC (total revenue to total population) with 95% point-wise confidence interval. Building age is estimated nonparametrically using a thin plate spline and interacted with the TperC (logged and demeaned), an indicator if the unit is detached, and the floor area (logged and demeaned). All smooth terms are significant at 1% with p-values approximately less than 0.001.

Figure 1.5: Nonparametric Building Age Interaction with LTLiability



This figure presents results from the estimation of (3.17) controlling for LTLiability (total long-term debt outstanding to total interest on debt) with 95% point-wise confidence interval. Building age is estimated nonparametrically using a thin plate spline and interacted with the LTLIability (logged and demeaned), an indicator if the unit is detached, and the floor area (logged and demeaned). All smooth terms are significant at 1% with p-values approximately less than 0.001.

Chapter 2

Benchmarking the Performance of U.S. Municipalities with Paul W. Wilson

Municipal governments provide varying bundles of goods, services and amenities for residents, who in turn are free to choose among municipalities and hence the varying offerings. In addition to police and fire protection, municipal governments may provide roads and streets, traffic management, trash collection, street cleaning, water services, libraries, and other services. In principle, municipalities compete with each other both in terms of taxation as well as provision of services. Grosskopf et al. (2001) note that competition among municipalities may create incentives to provide services efficiently by influencing citizens' willingness to pay for public services or their inclination to remain in the jurisdiction. Hayes et al. (1998), Grossman et al. (1999) and others find evidence that competition among local governments tends to enhance efficiency. At the same time, friction caused by real estate transaction fees, costs of commuting and job search and other factors may reduce competition among municipalities, perhaps leading to inefficient provision of services.

Grosskopf et al. (2001) also suggest that monitoring by voters may encourage efficiency among municipalities. Government officials may increase their probability of remaining in office by running local governments efficiently, particularly where residents are personally affected by local policy. Davis and Hayes (1993) argue that citizens are more likely to closely monitor local governments where taxes are high, and that housing owners are more likely to watch closely than renters because they have a larger stake in outcomes. Davis and Hayes (1993), Grossman and West (1994), Hayes and Wood (1995) and Hayes et al. (1998) find evidence that increased monitoring (proxied by tax rates or the degree of centralization by local governments) is associated with smaller or more efficient local governments. Here again, however, the effect of voter monitoring on local government efficiency may be limited by lack of accessibility to government officials, incumbents' inherent advantages over challengers in the polls, lack of information and other factors.

This paper examines the technical efficiency of U.S. municipal governments during 1997, 2002, 2007 and 2012 using data from the U.S. Census of Governments and the Annual Survey of State and Local Finances. The data cover the recent financial crisis of 2007–2008 and beyond when many municipalities struggled to meet expenses in the face of falling tax revenues. Several U.S. cities have filed for Chapter 9 bankruptcy in the aftermath of the financial crisis, including Vallejo, CA in 2008; Harrisburg, PA in 2011; Central Falls, RI in 2011; Stockton, CA in 2012; San Bernardino, CA in 2012; Detroit, MI in 2013; and Hillview, KY in 2015.[1] With bankruptcies, rising pension burdens, mandatory balanced-budget constraints, and limited and dwindling sources of future revenue, the question of whether municipal governments use scarce resources to provide services efficiently is an important topic

[1] The Harrisburg case was subsequently dismissed.

in the post-recession climate.

The idea that municipal governments should—but sometimes do not—provide services efficiently is an old one. Tiebout (1956) was the first to discuss competition in local government production, remarking that households sort based on their preferences for public goods and services. He notes that municipal competition increases overall efficiency. These efficiency gains are often capitalized in local property values. By contrast, Bruère et al. (1912) state that "the efficiency movement in cities grew out of recognition of the dependence of community welfare upon government activity" beginning in 1906, and that the efficiency movement "aims to remove city government from its isolation, and to make it the customary and accepted common agency for 'getting things done'." However, municipal governments provide a classic example of the principal-agent problem. Residents pay taxes and consume services, but must delegate management of municipal governments to politicians, bureaucrats and functionaries.

A number of empirical studies have examined efficiency among local governments (e.g., see the reviews by Tang, 1997; De Borger and Kerstens, 2000; Afonso, 2008; Da Cruz and Marques, 2014; De Oliveira Junqueira, 2015; and Narbón-Perpiñá and De Witte, 2018). The empirical analyses can be broadly divided into those that employ fully parametric methods along the lines of Aigner et al. (1977) and Meeusen and van den Broeck (1977) versus those that use fully nonparametric methods such as the data envelopment analysis (DEA) estimators proposed by Farrell (1957) and popularized by Banker et al. (1984) or free disposal hull (FDH) estimators proposed by Deprins et al. (1984). Both parametric and nonparametric studies have typically reported only point estimates of efficiency, with no inference. In studies of local government efficiency, nonparametric methods are more often used than parametric methods, and among studies using nonparametric methods, DEA estimators that impose convexity on the production set are used much more often than FDH estimators that do not impose convexity.[2] The choice between FDH and DEA estimators is not innocuous—DEA estimators are not statistically consistent if the production set is not convex, while FDH estimators are consistent regardless of whether the production set is convex (see Simar and Wilson, 2013, 2015 for details and discussion).

This study employs nonparametric methods, thereby avoiding troublesome functional form assumptions. The translog specification is often used for production and cost functions in parametric applications. Among the 31 papers listed by Narbón-Perpiñá and De Witte (2018, Table A2) that employ parametric methods to assess local government performance, 27 use a translog specification, Nikolov and Hrovatin (2013) and Pacheco, Sanchez, and Villena (2014) use a Cobb-Douglas specification (which of course is nested by the translog specification), and Hayes and Chang (1990) and Revelli (2009) use alternative model specifications.

However, municipal governments vary widely in terms of size, and several studies have noted that the parameters of a translog function are unlikely to be stable when the function is fit globally across units of widely varying size. See, for example, Guilkey et al. (1983) and Chalfant and Gallant (1985) for Monte Carlo evidence, and Cooper and McLaren (1996) and Banks et al. (1997) for empirical evidence involving consumer demand, Wilson and Carey (2004) for empirical evidence involving hospitals, and McAllister and McManus (1993), Mitchell and Onvural (1996), and Wheelock and Wilson (2001, 2012, 2018) for empirical evidence involving banks.

In addition, recent theoretical results on the properties of nonparametric efficiency estimators are used to test convexity versus non-convexity of municipal govern-

[2]The review by Narbón-Perpiñá and De Witte (2018, Table A2) lists 97 empirical studies of local government efficiency. Sixty-six of these studies used nonparametric estimators, while only 31 use parametric methods. Among the 66 papers employing nonparametric estimators, 50 use DEA estimators, 14 use FDH estimators, and 2 use both.

ments' production sets, and to test for differences in mean efficiency across regions as well as across time. The data reject convexity in favor of non-convexity, and this result casts substantial doubt on previous studies that have arbitrarily imposed convexity without testing.[3] This study, to our knowledge, is the first to combine nonparametric estimation of efficiency with rigorous, statistically valid inference-making to look at the efficiency of municipal governments.

The paper proceeds as follows. A statistical model and relevant estimators are presented in the next section. In addition, various statistical results needed for testing hypotheses about model features are also discussed in Section 2.1. The data used for estimation and inference are discussed in Section 2.2, and empirical results are presented in Section 2.3. Summary and conclusions are given in Section 2.4.

2.1 Methods for Estimation and Inference

To establish notation, let $X \in \mathbb{R}^p_+$ and $Y \in \mathbb{R}^q_+$ denote random vectors of input and output quantities, and similarly let $x \in \mathbb{R}^p_+$ and $y \in \mathbb{R}^q_+$ denote corresponding fixed, nonstochastic vectors of input and output quantities. The production set is the set of feasible combinations of input and output quantities, i.e.,

$$\Psi := \{(x, y) \mid x \text{ can produce } y\}, \tag{1.1}$$

[3] In addition to the 50 papers listed by Narbón-Perpiñá and De Witte (2018, Table A2) that employ DEA (but not FDH) to examine cities' efficiencies, see also and Charnes et al. (1989), Duncombe et al. (1997), Hájková and Hájek (2014). A search on Google Scholar on 18 December 2019 using the keywords "DEA", "efficiency" and "cities" yielded approximately 39,300 hits. A few studies (e.g., De Borger and Kerstens (1996), Balaguer-Coll et al. (2007a) and Geys and Moesen (2009) estimate cities' efficiencies using both FDH and DEA estimators and note that the estimates are different, but do not test whether the production set is convex (and hence whether consistency holds for DEA estimators).

which gives the set of possible inputs and outputs. Standard assumptions in the microeconomic theory of the firm include the following (e.g., see Shephard or Färe).

Assumption 2.1.1 Ψ *is closed.*

Assumption 2.1.2 *Production requires the use of some inputs, i.e.,* $(x, y) \notin \Psi$ *if* $x = 0,\ y \geq 0,\ y \neq 0$.

Assumption 2.1.3 *Inputs and outputs are strongly disposable, i.e.,* $\forall\,(x, y) \in \Psi$, *(i)* $\widetilde{x} \geq x \Rightarrow (\widetilde{x}, y) \in \Psi$ *and (ii)* $\widetilde{y} \leq y \Rightarrow (x, \widetilde{y}) \in \Psi$.

Assumption 2.1.1 ensures that the *efficient frontier* or *technology*

$$\Psi^{\partial} := \left\{(x, y) \mid (x, y) \in \Psi,\ (\gamma^{-1}x, \gamma y) \notin \Psi \text{ for any } \gamma \in (1, \infty)\right\} \tag{1.2}$$

consisting of the set of extreme points of Ψ is contained in Ψ. We define inequalities involving vectors on an element-by-element basis. Assumption 2.1.2 requires use of some inputs to produce any output greater than zero, thereby ruling out the existence of free lunches (throughout, inequalities involving vectors are defined on an element-by-element basis). Assumption 2.1.3 amounts to imposition of weak monotonicity on the frontier.

The frontier Ψ^{∂} provides a benchmark against which production units' performance can be measured. Units operating on the frontier are said to be *technically efficient*, while those operating under the frontier, in the interior of Ψ, are said to be *technically inefficient*. Several measures of technical (in)efficiency are employed in the literature. The Farrell (1957) input efficiency measure

$$\theta(x, y \mid \Psi) := \inf\left\{\theta \mid (\theta x, y) \in \Psi\right\} \tag{1.3}$$

gives the proportion by which input levels can be feasibly reduced without reducing output levels. Alternatively, the Farrell (1957) output efficiency measure

$$\lambda(x, y \mid \Psi) := \sup \{\lambda \mid (x, \lambda y) \in \Psi\}. \tag{1.4}$$

gives the feasible proportion by which output levels can be increased without increasing input quantities. Alternatively, the hyperbolic measure

$$\gamma(x, y \mid \Psi) := \inf \left\{\gamma > 0 \mid (\gamma x, \gamma^{-1} y) \in \Psi\right\} \tag{1.5}$$

proposed by Färe et al. (1985) gives the proportion by which output quantities can be increased while simultaneously reducing input quantities by the same proportion. Clearly, $\theta(x, y \mid \Psi) \leq 1$, $\lambda(x, y \mid \Psi) \geq 1$, and $\gamma(x, y \mid \Psi) \leq 1$ for all $(x, y) \in \Psi$. Moreover, due to Assumptions 2.1.1 and 2.1.3, $\gamma(x, y \mid \Psi) = 1$ for any $(x, y) \in \Psi^\partial$.[4]

It is important to note that Ψ, and hence Ψ^∂ as well as the measures $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ are unobserved. Consequently, they must be estimated from a sample $\mathcal{S}_n = \{(X_i, Y_i)\}_{i=1}^n$ of observed input-output pairs. There are several possibilities.

Deprins et al. (1984) estimate Ψ by the free disposal hull of the sample observations in $\mathcal{S}_n$, i.e.,

$$\widehat{\Psi}_{\text{FDH},n} := \bigcup_{(X_i, Y_i) \in \mathcal{S}_n} \left\{(x, y) \in \mathbb{R}_+^{p+q} \mid x \geq X_i,\ y \leq Y_i\right\}. \tag{1.6}$$

This estimator does not impose convexity on Ψ. Alternatively, Ψ can be estimated using the VRS (DEA) proposed by Farrell (1957) and popularized by Banker et al.

[4]In principle, for some $(x, y) \in \Psi^\partial$ one might have $\theta(x, y \mid \Psi) < 1$ or $\lambda(x, y \mid \Psi) > 1$ if the frontier is parallel to either all of the input axes or all of the output axes in some regions. However, this is ruled out by additional assumptions made below to define a statistical model.

(1984). The VRS (DEA) estimator of Ψ amounts to the convex hull of $\widehat{\Psi}_{\text{FDH}}$, and is given by

$$\widehat{\Psi}_{\text{VRS},n} := \left\{(x, y) \in \mathbb{R}^{p+q} \mid y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{1.7}$$

where $\boldsymbol{X} = \begin{pmatrix} X_1, \ \ldots, \ X_n \end{pmatrix}$ and $\boldsymbol{Y} = \begin{pmatrix} Y_1, \ \ldots, \ Y_n \end{pmatrix}$ are $(p \times n)$ and $(q \times n)$ matrices of input and output vectors, respectively; $\boldsymbol{i}_n$ is an $(n \times 1)$ vector of ones, and $\boldsymbol{\omega}$ is a $(n \times 1)$ vector of weights. The estimator $\widehat{\Psi}_{\text{VRS},n}$ imposes convexity, but allows for VRS. Dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (1.7) results in the constant returns-to-scale (CRS) estimator $\widehat{\Psi}_{\text{CRS},n}$ of Ψ. Note that CRS holds if and only if $(\alpha x, \alpha y) \in \Psi$ for all $(x, y) \in \Psi$ and $\alpha \in (0, \infty)$.

Estimators of the efficiency measures $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ defined above are obtained by substituting either $\widehat{\Psi}_{\text{FDH},n}$, $\widehat{\Psi}_{\text{VRS},n}$ or $\widehat{\Psi}_{\text{CRS},n}$ for Ψ in (1.3)–(1.5) (respectively). For example, substituting $\widehat{\Psi}_{\text{VRS},n}$ for Ψ in (1.3)–(1.5) yields

$$\widehat{\theta}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \min_{\theta, \boldsymbol{\omega}} \left\{\theta \mid y \leq \boldsymbol{Y}\boldsymbol{\omega},\ \theta x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{1.8}$$

$$\widehat{\lambda}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \max_{\lambda, \boldsymbol{\omega}} \left\{\lambda \mid \lambda y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\} \tag{1.9}$$

and

$$\widehat{\gamma}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \min_{\gamma, \boldsymbol{\omega}} \left\{\gamma \mid \gamma^{-1} y \leq \boldsymbol{Y}\boldsymbol{\omega},\ \gamma x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}. \tag{1.10}$$

The estimators in (1.8)–(1.9) can be computed using linear programming methods. The hyperbolic estimator in 1.10 is a non-linear program and can be computed easily using the algorithm developed by Wilson (2011).

Substituting $\widehat{\Psi}_{\text{FDH},n}$ into (1.3)–(1.5) results in integer programming problems. Nonetheless, the resulting estimators can be computed easily. In particular, let $\mathcal{D}_{x,y}$ denote the set indices of points in $\mathcal{S}_n$ dominating (x, y), i.e., $\mathcal{D}_{x,y} = \{i \mid (X_i, Y_i) \in \mathcal{S}_n,\ X_i \leq x,\ Y_i \geq y\}$. Then

$$\widehat{\theta}_{\text{FDH}}(x, y \mid \mathcal{S}_n) = \min_{i \in \mathcal{D}_{x,y}} \max_{j=1,\ldots,p} \left(\frac{X_i^j}{x^j} \right), \tag{1.11}$$

where for a vector a, a^j denotes its j-th component. The output-oriented estimator can be computed by solving

$$\lambda_{\text{FDH}}(x, y \mid \mathcal{S}_n) = \max_{i \in \mathcal{D}(x,y)} \min_{j=1,\ldots,q} \left(\frac{Y_i^j}{y^j} \right), \tag{1.12}$$

and the hyperbolic estimator can be computed by solving

$$\widehat{\gamma}_{\text{FDH}}(x, y \mid \mathcal{S}_n) = \min_{i=1,\,\ldots,\,n} \left(\max_{\substack{j=1,\,\ldots,\,p \\ k=1,\,\ldots,\,q}} \left(\frac{x_i^j}{x^j}, \ \frac{y^k}{y_i^k} \right) \right) \tag{1.13}$$

as shown by Wilson (2011).

Of course, in addition to estimation, inference is needed before anything can be learned from data, and inference requires a well-defined statistical model. The following assumptions, together with Assumptions 2.1.1–2.1.3, complete the model and are sufficient to establish properties of the estimators that will be used later.

Assumption 2.1.4 *(i) The random variables (X, Y) possess a joint density f with support $\mathcal{D} \subset \Psi$; and (ii) f is continuously differentiable on $\mathcal{D}$.*

Assumption 2.1.5 *(i) $\mathcal{D}^* := \{\theta(x, y \mid \Psi)x,\ y) \mid (x, y) \in \mathcal{D}\} = \{(x,\ \lambda(x, y \mid \Psi)y) \mid (x, y) \in \mathcal{D}\} = \{(\gamma(x, y \mid \Psi)x,\ \gamma(x, y \mid \Psi)^{-1}y) \mid (x, y) \in \mathcal{D}\} \subset \mathcal{D}$; (ii) $\mathcal{D}^*$ is compact; and (iii) $f(\theta(x, y \mid \Psi)x, y) > 0$ for all $(x, y) \in \mathcal{D}$.*

Assumption 2.1.4–2.1.5 describe a probability model (which is a component of every statistical model). Assumption 2.1.4(i) means that any observed input-output data fall within the production set Ψ, while Assumption 2.1.4(ii) imposes some smoothness on the joint density f of inputs and outputs. Assumption 2.1.5 implies that the density f is strictly positive along any part of the frontier where observations in $\mathcal{D}$ might be projected to by one of the efficiency measures introduced above. In addition, Assumption 2.1.5 rules out infinite input levels, which should not be troubling and which simplifies derivation of the estimators' properties.

The VRS estimators described above require the next two assumptions.

Assumption 2.1.6 *$\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ are three times continuously differentiable on $\mathcal{D}$.*

Assumption 2.1.7 *$\mathcal{D}$ is* almost strictly convex*; i.e., for any $(x, y), (\widetilde{x}, \widetilde{y}) \in \mathcal{D}$ with $(\frac{x}{\|x\|}, y) \neq (\frac{\widetilde{x}}{\|\widetilde{x}\|}, \widetilde{y})$, the set $\{(x^*, y^*) \mid (x^*, y^*) = (x, y) + \alpha((\widetilde{x}, \widetilde{y}) - (x, y))$ for some $0 < \alpha < 1\}$ is a subset of the* interior *of $\mathcal{D}$.*

Kneip et al. (2008) require only two-times differentiability to establish the existence of a limiting distribution for VRS estimators, but the stronger Assumption 2.1.6 is needed by Kneip et al. (2015) to establish results on moments of the VRS estimators. Assumption 2.1.6 imposes more smoothness on the frontier than required by Kneip et al. (2008). Recalling that the strong (i.e., free) disposability assumed in Assumption 2.1.3 implies that the frontier is weakly monotone, Assumption 2.1.7 strengthens this by requiring the frontier to be strictly monotone with no constant segments. This is also needed to establish properties of moments of the VRS estimators.

Alternatively, when FDH estimators are used, Assumptions 2.1.6 and 2.1.7 can be replaced by the following assumption.

Assumption 2.1.8 *(i) $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ are twice continuously differentiable on $\mathcal{D}$; and (ii) all the first-order partial derivatives of $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ with respect to x and y are nonzero at any point $(x, y) \in \mathcal{D}$.*

Assumption 2.1.8 strengthens the assumption of strong disposability in 2.1.3 by requiring that the frontier is strictly monotone and does not possess constant segments.[5] Finally, part (i) of Assumption 2.1.8 is weaker than Assumption 2.1.6; here the frontier is required to be smooth, but not as smooth as required by Assumption 2.1.6.[6] Assumptions 2.1.1–2.1.5 and Assumption 2.1.8 comprise a statistical model appropriate for use of FDH estimators of technical efficiency, while Assumptions 2.1.1–2.1.7 comprise a statistical model appropriate for use of VRS estimators of technical efficiency.[7]

Under the appropriate set of assumptions—either Assumptions 2.1.1–2.1.5 and Assumption 2.1.8 for the FDH estimators, or Assumptions 2.1.1–2.1.7 for the VRS estimators—the various efficiency estimators have well-developed statistical properties. Kneip et al. (1998) establish consistency and the rate of convergence of the input-oriented VRS efficiency estimator. Kneip et al. (2008) derive its limiting distribution. Park et al. (2000) and Daouia et al. (2017) provide the rate of convergence and the limiting distribution of the input-oriented FDH efficiency estimator. These results extend trivially to the output orientation. Wheelock and Wilson (2008) extend the results to the hyperbolic FDH efficiency estimator, and Wilson (2011) establishes similar properties for the hyperbolic VRS estimator. In all cases, the FDH estimators converge at rate $n^{1/(p+q)}$, while the VRS estimators converge at the faster

[5]Constant segments, neither increasing nor decreasing in inputs, might occur if outputs are discrete (i.e., "lumpy"), rather than continuous.

[6]Assumption 2.1.8 is slightly stronger, but much simpler than assumptions AII–AIII in Park et al. (2000).

[7]Additional, stronger assumptions are needed for constant returns-to-scale efficiency estimators. See Kneip et al. (2015) for additional discussion.

rate $n^{2/(p+q+1)}$. Consequently, both FDH and VRS estimators suffer from the "curse of dimensionality" that typically plagues nonparametric estimators. Wilson (2018) provides evidence that the problem can be mitigated in many situations by using dimension-reduction techniques.

By construction, both the FDH and VRS efficiency estimators are biased. This is due to the fact that $\widehat{\Psi}_{\text{FDH},n} \subseteq \Psi$ and $\widehat{\Psi}_{\text{VRS},n} \subseteq \Psi$ provided Ψ is convex. Kneip et al. (2015) provide results on moments of the input-oriented efficiency estimators, and the results extend trivially to the output-oriented estimators. Kneip et al. (2018) extend these results to the hyperbolic VRS efficiency estimator, and Wilson (2019) extends the results toe the hyperbolic FDH estimator. In all cases, the bias of the estimators disappears at the same rate at which the estimators converge. Consequently, for a convergence rate of n^{κ}—where $\kappa = 1/(p+q)$ for the FDH estimators, $\kappa = 2/(p+q+1)$ for the VRS estimators, or $\kappa = 2/(p+q)$ for the CRS estimators—standard central limit theorem (CLT) results (e.g., the Lindeberg-Feller CLT) do not hold for mean efficiency unless $\kappa > 1/2$. In the CRS case, this means that the usual CLTs hold only if $(p+q) < 4$. In the VRS case, the usual CLTs hold only if $(p+q) < 3$. In the FDH case, standard CLTs hold only if $p+q < 2$.[8]

In addition to deriving properties on moments of the input-oriented efficiency estimators, Kneip et al. (2015) provide new CLTs (for all values of κ) for means of the input-oriented FDH, VRS and CRS estimators. Kneip et al. (2016) use these results to establish asymptotically normal test statistics that can be used to test convexity versus non-convexity of Ψ, CRS versus VRS (provided Ψ is convex), and differences in mean efficiency across groups of producers. All these results extend trivially to the output-oriented cases. Kneip et al. (2018) extend these results to the hyperbolic VRS

[8]In other words, standard CLT results hold in the FDH case if and only if $p = 1$ and output is fixed and constant, or $q = 1$ and input is fixed and constant.

estimator, and Wilson (2019) extends the results to the hyperbolic FDH estimator.

2.2 Data and Variable Specification

We use data from the the Annual Survey of State and Local Government Finances, the Annual U.S. Building and Permit Survey, the U.S. Census of Governments, the U.S. Bureau of Labor Statistics (BLS) and the Federal Bureau of Investigation (FBI) to define input and output variables. Our variable specifications are broadly similar to those use in other studies of local governments' efficiency levels. All dollar amounts are measured in terms of thousands of constant, 2010 U.S. dollars.

The majority of local-government efficiency studies employ a single input variable to account for resources used to produce goods and services. Total current operating expenditures is the most widely specified input; examples include Štastná and Gregor (2015) and Radulović and Dragutinović (2015). Alternatively, a few studies have specified input as total expenditures (e.g., Hayes and Chang, 1990) or financial expenditures (e.g., De Borger and Kerstens, 1996). We adopt the former, more typical approach and specify a single input (denoted by X, and hence $p = 1$) consisting of total current operating expenditures reported in the Annual Survey of State and Local Government Finances. This survey is the only source of comprehensive, nationwide data on local government finances. Unfortunately, the survey only provides a full sample of local governments every five years (specifically, in years ending in '2' or '7').

In recent years, local governments have seen rising costs, exceeding the rate of increasing costs in the private sector, which is likely reflected in our input measure. Berry and Lowery (1984) speculate that Baumol's "cost disease" may be driving this difference, since many local public goods are labor-intensive or hand-produced.

Moreover, there has been an increasing level of concentration of government provision of goods and services at the federal level. Baicker et al. (2012) suggest this may be due to the growing importance of certain budget components including education, health, and welfare programs. By contrast, total and financial expenditures (as opposed to operating expenditures, which we use) include expenditures on total capital outlays, which is susceptible to volatility due to the nature of government spending. In this sense, our variable captures short-run operating costs.

We specify $q = 6$ output variables, including total population (Y_1), total charges for sewerage and waste management (Y_2), the reciprocal of the total crime rate (Y_3), total land area (in square miles) (Y_4), total building permits (Y_5) and and the employment rate (Y_6). Our specification of outputs reflects the wide variety of goods and services provided by municipal governments. Total population is one of the most frequently specified outputs in the literature on local-government efficiency (see lo Storto, 2013 and Athanassopoulos and Triantis, 1998) and serves as a proxy the scope of demand for publicly provided goods and services. We use total charges for sewerage and waste removal or treatment to account for communal service administration. Worthington (2000) utilizes a similar output measure related to the municipal sewerage system and waste collection, while Balaguer-Coll et al. (2007b) use the number of street lights.[9] Data for Y_1 and Y_2 are obtained from the Annual Survey of State and Local Government Finance.

We use the reciprocal of the total municipal crime rate, Y_3, to capture the degree of public safety provided through law enforcement services. These data are obtained from the FBI's Uniform Crime Reporting Statistics which are voluntarily reported by police departments at the local level. The FBI defines total violent

[9]Alternative measures of communal service administrative include number of highway miles and square footage of green space. Unfortunately, these data are not readily available for U.S. municipalities.

crimes to include murders and non-negligent manslaughter, legacy rape, revised rape, robbery, and aggravated assault. Total property (i.e., "non-violent") crimes include offenses such as burglary, larceny, and motor-vehicle theft. The total crime rate includes both total violent as well as total property crimes. We employ the reciprocal of the total crime rate so that our measure of safety, Y_3, increases as crime decreases.

We use total land area (in square miles), Y_4, as a measure of public services provided, following Grossman et al. (1999), obtained from the U.S. Census of Governments. Larger land areas require more infrastructure and public good servicing, such as highway repair and sewerage connections. The total number of unit-level building permits issued in a given year, Y_5, provides a measure of the amount of administrative services provided.[10] Data on building permits are obtained from the Annual U.S. Building and Permit Survey and are summed over months to obtain annual figures. Finally, data on the annual employment rate, Y_6, are obtained from the BLS Local Area Unemployment Statistics. All of our data are at the level of municipal governments in a given year.[11] After eliminating observations with missing values for one or more of our variables, we have 649, 730, 746 and 800 observations for 1997, 2002, 2007 and 2012 (respectively), for a total of 2,925 observations.

We assume that all municipalities operate in the same production set Ψ defined by (1.1), and consequently face the same frontier Ψ^{∂} in the seven-dimensional input-output space. Note, however, that local governments may have very different scales and budget plans, and hence may operate in different regions of the production set or under different parts of the frontier. The model described in Section 2.1 is fully non-parametric, and hence quite flexible. The assumptions listed in Section 2.1 impose

[10] We use the number of permits issued for individual units, rather than number of buildings for which permits have been issued. In dense urban environments, one might observe multiple units (e.g., condominium units) in a single building for which building permits have been issued.

[11] We use the U.S. Census Bureau's definition of local government, corresponding to type code 2 in the U.S. Census of Governments' 14-digit government ID code.

only minimal restrictions involving free-disposability, continuity and some smoothness of the frontier. Note that there is no assumption of convexity of Ψ, which we test below in Section 2.3.

Although our non-parametric model is highly flexible, there is a price to pay in terms of the well-known "curse of dimensionality." Wilson (2018) discusses dimension-reduction in the context of nonparametric efficiency estimation, and presents diagnostics to indicate whether reducing dimensionality might be advantageous. As discussed in Section 2.1, the FDH, VRS and CRS estimators converge at rate n^{κ}, where $\kappa = 1/(p+q)$ for FDH estimators, $\kappa = 2/(p+q+1)$ for VRS estimators and $\kappa = 2/(p+q)$ for CRS estimators. With the $(p+q) = 7$ dimensional specification described above, the convergence rates are $n^{1/7}$, $n^{1/4}$ and $n^{2/7}$ for FDH, VRS and CRS estimators, respectively. Moreover, as noted above, the number of observations in each period range from 649 to 800. The *effective parametric sample size* defined by Wilson (2018) is then, in the worst case, $649^{1/7} \approx 6$ for FDH estimators, $649^{1/4} \approx 25$ for VRS estimators and $649^{2/7} \approx 40$ for VRS estimators. In other words, with a sample size of 649, FDH estimators should be expected to result in estimation error of the same order one would achieve with a typical parametric estimator (converging at the root-n rate) and only 6 observations. With VRS (or CRS) estimators, one should expect estimation error of the same order that 25 (or 40) observations would provide in a parametric model. Of course, consistency of the VRS estimators requires convexity of Ψ, and consistency of the CRS estimators requires in addition CRS. It remains to be seen whether Ψ satisfies such restrictions.[12]

Wilson (2018) also suggests examining the ratio R_y of the largest eigenvalue of

[12] Of course, the notion of *effective parametric sample size* defined by Wilson (2018) presupposes that one has a correctly specified parametric model. As Robinson (1988) notes, the root-n parametric convergence rate means that estimators converge quickly to the wrong thing in a mis-specified model; Robinson-1988 refers to this as root-n *in*consistency.

the moment matrix $\boldsymbol{YY'}$ to the sum of eigenvalues for $\boldsymbol{YY'}$. moment matrices. Our data yield values 99.38, 99.75, 99.86 and 99.64 for R_y in 1997, 2002, 2007 and 2012. If one considers a set of rays from the origin passing through each observation in the six-dimensional output space $\mathbb{R}^6_+$ for each year, the values of R_y indicate that these rays lie in a very tight bundle and are very similar in terms of their angles with respect to each axis. The results also indicate that the data contain almost no information about marginal rates of transformation between outputs. The smallest of these values, 99.38, is well above the level needed for dimension reduction to likely reduce mean square error of either DEA or FDH estimates as indicated by the simulation results reported by Wilson (2018). Consequently, we compute the $(1 \times n)$ principal component vector $Y_* = E_y' \boldsymbol{Y}$ where E_y is the $(q \times 1)$ eigenvector corresponding to the largest eigenvalues of the moment matrix $\boldsymbol{YY'}$. Given the values for R_y listed above, it is clear that these principle components contain almost all of the independent information in the $q = 6$ outputs specified above. Except as noted below, all estimation is done using the single input variable X and the (output) principal component Y_*. In this two-dimensional setting, the convergence rates of the FDH, VRS and CRS estimators are $n^{1/2}$, $n^{2/3}$ and n^1, respectively. The simulation results of Wilson (2018) provide clear evidence that relying only on X and Y_* for estimation likely results in less estimation error than would be the case with five dimensions. This is true regardless of whether the technology is homothetic, contrary to what is suggested by Färe and Lovell (1988) and Olesen and Petersen (2016).

Summary statistics for our input and output variables as well as the principal component Y_* are shown in Table 2.1. For each variable, the table shows the minimum value, first quartile (Q1), median, mean, third quartile (Q3) and the maximum value. The wide range of city sizes and our use of data spanning 15 years results in substantial variation in each of the variables as reflected in the table. Comparing differences between the median and Q1 and between Q3 and the median for the input and output variables reveals that the marginal distributions are skewed to the right.

2.3 Empirical Results

As a preliminary step, we further investigate the efficacy of dimension reduction, we first estimate the hyperbolic efficiency in (1.5) for municipalities in each year using the full-dimensional data with one input and six outputs. We then repeat this exercise using the input variable X and the reduced-dimensional output variable Y_*. In both cases, we obtain estimates from the FDH, VRS and CRS estimators described above in Section 2.1. Table 2.2 shows the number of observations in each year as well as counts of the number of estimates equal to one in each of the resulting six scenarios. As discussed by Wilson (2018), large proportions of efficiency estimates equal to one, especially among FDH estimates, may indicate the need for dimension reduction.

Each of the FDH, VRS and CRS estimators are biased, but the bias is largest for the FDH estimator, and smallest for the CRS estimator. The counts in Table 2.2 where the full data are used reveal that about half of the FDH estimates are equal to one in each year. By contrast, a much smaller proportion of the VRS estimates equal one, and less than 1.5 percent of the CRS estimates are equal to one in each year. Taken together, the results indicate that most of the inefficiency that one would find using either the VRS or CRS estimators is merely an artifact of the convexity imposed by both estimators, and in addition the assumption of CRS imposed by the CRS estimator. As such, the estimates in Table 2.2 obtained with the full data, in particular the large proportion of estimates equal to one when the FDH estimator is used, make clear the need for dimension reduction. The evidence is clear that there are too many dimensions for the sample size (see Wilson, 2018 for discussion).

Furthermore, the differences in proportions obtained with the FDH estimator and those obtained with the VRS estimator suggest that Ψ may not be convex. This is investigated further below.

Table 2.2 also shows counts of the numbers of observations with efficiency estimates equal to one when the reduced-dimensional data are used. In the last three columns of Table 2.2, the counts for the FDH estimator are larger than the counts for the VRS estimator, which in turn are larger than the counts for the CRS estimator. This is to be expected since the bias decreases as one moves from the FDH to the VRS and then the CRS estimators. However, the counts for the FDH estimator with the reduced-dimensional data amount to about 10 percent of the counts for the FDH estimator with the full data. Overall, the results in Table 2.2 provide evidence (in addition to the values of R_y and the effective parametric sample sizes discussed earlier in Section 2.2) that dimension reduction likely reduces estimation error relative to what would be obtained working in the full, seven-dimensional space. Consequently, we employ dimension reduction and work in the two-dimensional space of the variables X and Y_* for the remainder of the paper.

Having specified the fully nonparametric model presented in Section 2.1, we are confronted with three different estimators. In terms of restrictiveness, the FDH estimator is the least restrictive, followed by the VRS estimator which imposes convexity of the production set Ψ, and finally by the CRS estimator which imposes CRS in addition to convexity. Moreover, there is a tradeoff between the rate of convergence and the restrictiveness of the estimators; i.e., the FDH estimator converges slower than the other two, while the CRS estimator has the fastest rate.[13] In many applied studies in the literature, the choice between FDH, VRS and CRS estimators

[13]If CRS holds, then the VRS estimator attains the faster rate of the CRS estimator, i.e., n^{κ} with $\kappa = 2/(p+q)$ as proved by Kneip et al. (2016).

often appears to be made arbitrarily, or worse, perhaps to avoid excessive numbers of estimates equal to one. As the results for the full data (or even for the reduced data) in Table 2.2 indicate, the VRS and CRS estimators produce substantially fewer estimates equal to one than the FDH estimator. One might suspect that this explains why the VRS and CRS estimators are used with much greater frequency than the FDH estimator. However, whether the VRS estimator is appropriate depends on whether the production set is convex. If Ψ is convex, whether the CRS estimator is appropriate depends further on whether CRS holds.

To decide whether the FDH or VRS estimator should be used, we first test the null hypothesis of convexity of Ψ versus the alternative hypothesis that Ψ is not convex. We use the test developed by Kneip et al. (2016), augmented by results from Simar and Wilson (2019). The test described by Kneip et al. (2016) involves randomly splitting the sample for a given year into two independent subsamples of sizes n_1 and $n_2 = n - n_1$ and comparing the sample mean of VRS efficiency estimates from the first subsample with n_1 observations and the sample mean of FDH efficiency estimates from the second subsample with n_2 observations. The null hypothesis of convexity is rejected when the difference between the two sample means is "large." The subsample sizes are chosen by setting $\widetilde{n}_1^{2/(p+q+1)} = \widetilde{n}_2^{1/(p+q)}$ and $\widetilde{n}_1 + \widetilde{n}_2 = n$, solving for $\widetilde{n}_1$, and then setting $n_1 = \lfloor \widetilde{n}_1 \rceil$ and $n_2 = n - n_1$, where $\lfloor a \rceil$ denotes rounding of $a \in \mathbb{R}$ to the closest integer. The first subsample is used to compute VRS estimates, and the second is used to compute FDH estimates. The test statistic given in equation (50) of Kneip et al. (2016) involves the difference of the means of these two sets of estimates, with generalized jackknife estimates of biases and corresponding sample variances, and is asymptotically normally distributed with mean zero and unit variance. The test is a one-sided test since under the null the two means should be roughly similar, but the statistic should diverge from $N(0,1)$ with increasing departures from the null resulting in the mean of the FDH estimates exceeding the mean of the VRS estimates. The statistic given in equation (50) of Kneip et al. (2016) is defined in terms of input-oriented estimators, but the asymptotic normality result obtained by Kneip et al. (2016) extends trivially to output-oriented and hyperbolic estimators. In each case, statistics are defined so that “large” positive values indicate rejection of the null hypothesis.

Splitting the sample is necessary to maintain independence between the means of the FDH and VRS estimates as explained by Kneip et al. (2016). Although the test is valid for a single, random split, one may obtain different results by splitting the sample differently. In fact, under the null the test statistics are asymptotically distributed $N(0, 1)$, and hence by the probability integral transform, the corresponding p-values (which are also random variables) are distributed uniformly on the $(0, 1)$ interval. If one had m samples, each could be randomly split, then m test statistics and corresponding p-values could be computed, one for each sample. One could then use the sample mean of the test statistics, which under the null would be asymptotically distributed $N(0, m^{-1})$. Alternatively, one could use the Kolmogorov-Smirnov one-sample test to test whether the resulting p-values are uniformly distributed; if not, one should reject the null hypothesis of convexity of Ψ.

Given a single sample, splitting repeatedly results in dependence from one split to the next. Hence the limiting distribution of the sample mean of test statistics from m splits of a *single* sample is unknown. For the same reason, the usual tables for the Kolmogorov-Smirnov statistic are invalid and should not be used. Simar and Wilson (2019) develop a bootstrap method for estimating (i) p-values for tests based on either sample means of test statistics across m splits of a given sample (test #1), and (ii) the sampling distribution of the Kolmogorov-Smirnov statistic for the m p-values obtained from m splits of the same sample, permitting inference about convexity while removing the arbitrariness of a single random split of the sample. Simulation results provided by Simar and Wilson (2019) indicate that the proposed bootstrap method results in tests with good size and power properties.

For each of the four years covered by our data, we test convexity versus non-convexity of Ψ using the Kneip et al. (2016) test with 1,000 sample-splits and the bootstrap proposed by Simar and Wilson (2019) with 1,000 bootstrap replications. For each year we conduct three tests, using input-oriented, output-oriented, and hyperbolic FDH and VRS estimators. The resulting values of test statistics and p-values are shown in Table 2.3. Using the input-oriented or hyperbolic estimators, the largest p-value is 0.007. Using output-oriented estimators, convexity is not rejected using test #1 based on the sample mean of the Kneip et al. (2016) statistic over 1,000 sample-splits, nor is convexity rejected for 1997 and 2002 using test #2 based on the Kolmogorov-Smirnov statistic. But overall, the data provide ample evidence against convexity of Ψ, and consequently we rely on FDH estimators for the remainder of our analysis since both the VRS and CRS estimators require convexity for statistical consistency, whereas the FDH estimators do not require convexity.[14]

Table 2.4 reports summary statistics for FDH efficiency estimates in each year for each orientation (i.e., input, output and hyperbolic). To facilitate comparison, we report statistics on reciprocals of the output-oriented estimates, so that in each case larger values correspond with greater efficiency (i.e., less inefficiency). As discussed by Wilson (2011), the levels of inefficiency in the input and output orientations may differ due to curvature of the frontier Ψ^{∂} and where municipalities lie in the input-output space. This is apparent from the results in the first two panels of Table 2.4, where it is evident that the distribution of output-oriented estimates (after taking

[14] The simulation results of Wilson (2018) indicate that the FDH estimator often yields smaller mean square error than the VRS estimator after dimension reduction, even if the underlying production set is convex.

reciprocals) lies to the right of the distribution of the input-oriented estimates. Wilson (2011) also discusses how the hyperbolic measure of efficiency represents a compromise between the input and output orientations, and is less sensitive to curvature of the frontier. The results in Table 2.4 reveal that mean estimated efficiency is higher for the hyperbolic estimates than for either the input- or output-oriented estimates.

Regardless of the direction in which efficiency is measured, the overall, qualitative results in Table 2.4 are similar across the three orientations. Each of the three sets of estimates suggest that technical efficiency declined from 1997 to 2012. The input-oriented estimates show a tiny increase in mean estimated efficiency from 2002 to 2007, and the hyperbolic estimates show no change from 2002 to 2007, but otherwise, mean estimated efficiency declines from one period to the next in Table 2.4.

To examine whether these changes are statistically significant, we use the test described by Kneip et al. (2016, Section 3.1.1) to test for significant differences between the means reported in Table 2.4 from one period to the next. This test permits the frontier to either be the same or to be different across periods, and does not require random splitting of the sample as with the convexity test described above since the data naturally fall into two groups. Even with reduced dimensionality, the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold for means of the FDH efficiency estimates as discussed in Kneip et al. (2015, 2016). As with the convexity test discussed above, the test statistic given by equation (18) of Kneip et al. (2016) involves not only the difference in sample means of efficiency estimates in a pair of years, but also the corresponding difference in generalized jackknife estimates of bias. The test extends trivially to the output-orientation, and due to Theorem 3.1 of Wilson (2019), it also extends easily to the hyperbolic orientation. The results in Table 2.5 indicate that the test soundly rejects the null hypothesis of no change in mean efficiency in every case. For 1997–2002, 2002–2007 and 1997–2012, the test statistics are negative, indicating declines in mean efficiency. However, for 2007–2012 the test statistics are positive, indicating that mean efficiency increased after 2007. This is in contrast to the means reported earlier in Table 2.4 where the sample means of efficiency estimates decline from 2007 to 2012. But as the results of Kneip et al. (2015) make clear, sample means of FDH efficiency estimates involve bias, which is ignored in Table 2.4 but is accounted for in Table 2.5. Overall, the results in Table 2.5 indicate that mean efficiency declined from 1997 to 2007, and although there was improvement from 2007 to 2012, mean efficiency declined on net between 1997 and 2012.

Due to the regional heterogeneity among state and local governments, it is important to consider possible differences in mean efficiencies across regions. We divide municipal governments represented in our sample into the four regions defined by the U.S. Census Bureau: Northeast (1), Midwest (2), South (3), the West (4). These census regions are illustrated in Figure 2.1. Then for each of the four census regions, we test for differences in mean efficiency across pairs of years, analogous to the previous tests using the entire sample and ignoring regions. Results of these tests are reported in Table 2.6. While the results in Table 2.5 indicate decreases in mean efficiency from 1997 to 2002 and from 2002 to 2007, the results in Table 2.6 reveal that these results were not uniform over the different regions. The results in Table 2.6 indicate that mean efficiency increased over these periods in the Northeast, while in the Midwest and South, there is scant evidence of any change. In the west, mean efficiency decreased over each pair of years. For 2007 to 2012, the results suggest efficiency improved in the Northeast and South, but declined in the Midwest and West.

The results presented so far suggest changes in mean efficiency over time, and that these changes were not uniform over regions. To gain further insight, we examine pairs of years 1997–2002, 2002–2007, and 2007–2012 and apply the test of "separability" developed by Daraio et al. (2018), treating time as a binary "environmental" variable. The separability test in this case amounts to a test of whether time affects the frontier, i.e., whether the frontier shifts over time. The test requires splitting the randomly sorted, pooled sample of observations for each of two years, into two subsamples of sizes $n_1 = \lfloor n/2 \rfloor$ and $n_2 = n - n_1$, where $\lfloor a \rfloor$ denotes the largest integer not greater than $a \in \mathbb{R}$. Then mean efficiency estimates from the first subsample are compared with means of efficiency estimates from the second subsample, but in the second subsample, efficiency is estimated separately for each of the two years, in effect conditioning on time. Generalized jackknife estimates corresponding to each of the three sets of efficiency estimates are also computed. See Daraio et al. (2018) for details.

As with the convexity tests described above, the random splitting of the sample into two independent groups introduces noise and ambiguity, which can be removed using the bootstrap method of Simar and Wilson (2019) as discussed above in connection with the test for convexity of the production set. We again use 1,000 sample-splits, and 1,000 bootstrap replications. Results of the tests are shown in Table 2.7. The largest p-value (0.011) occurs with test #2 based on the Kolmogorov-Smirnov statistic. The data strongly reject separability with respect to time. Taken together with the results in Table 2.5, it is evident that while mean efficiency changed over time, the frontier also changed over time.

Table 2.8 presents results for similar tests of separability with respect to region for each of the four years represented in our sample. Here again, the data for a given year must be split into two subsamples of equal size. Efficiency is estimated over all observations in the first subsample, and independently over observations for each region in the second subsample. Then the mean of the unconditional estimates in the first subsample is compared with the mean of the conditional (on regions) estimates from the second subsample while accounting for bias using generalized jackknife estimates of bias. Again, see Daraio et al. (2018) for details. The bootstrap method of Simar and Wilson (2019) is used again, with 1,000 sample-splits and 1,000 bootstrap replications. We report results in Table 2.8 for tests based on sample means of test statistics across the 1,000 sample-splits (test #1) and tests based on the Kolmogorov-Smirnov statistic for the p-values corresponding to each of the 1,000 sample-splits (test #2). The largest p-value in Table 2.8 is 0.023, and all others are smaller than 0.000. Consequently, the data provide clear evidence against separability (i.e., a common frontier) across census regions.[15]

Having rejected a common frontier across regions in each year, we report in Table 2.9 sample means of hyperbolic FDH efficiency estimates by region and by year in order to provide an idea of the levels of inefficiency in each region and each year.[16] However, any comparisons of mean efficiency across years within a region should be made in conjunction with the results reported earlier in Table 2.6. These results, combined with those from the separability tests, suggest that municipalities in the different regions face different operating environments and constraints. The West region, in particular, has been the home of a number of large bankruptcies, including Stockton, CA and San Bernardino, CA.

[15] Clearly, one should not estimate efficiency over all municipalities in a given year, and then regress the efficiency estimates on dummy variables for the various regions and perhaps other environmental variables for the reasons given by Simar and Wilson (2007, 2011).

[16] Results for input- and output-oriented FDH efficiency estimates are similar qualitatively, but are omitted here to conserve space. The additional results are available from the authors on request.

2.4 Summary and Conclusions

Previous literature examining the municipal efficiency have failed to test for the convexity of Ψ, with the vast majority using VRS, DEA estimators. By testing for convexity versus non-convexity, we are able to let the data drive the choice of appropriate estimator. Our results suggest that Ψ is indeed non-convex in every year of our sample. Additionally, this is the first paper employing a fully nonparametric framework to gain insight on regional efficiency of U.S. municipalities. Since our results strongly rejected convexity, we estimated efficiency using FDH estimators, which are efficient in either case. To gain further insight, we also examine the technology and productivity of local governments. Since this is the first paper to test for this, we are unable to compare out results directly, but we are confident our results provide insight into efficiency of local governments and contributes significantly to the existing literature.

As previously described, by exploiting collinearity in the data to reduce dimensions to their first principle components, we went from a seven-dimensional problem to a two-dimensional problem. Results from Wilson (2018) indicate that this substantially reduces mean square error of efficiency estimators. Moreover, the simulation evidence provided by Wilson (2018) suggests that when production sets are convex, FDH estimates often have less mean square error than DEA estimators after dimension reduction.

Our results suggest that municipalities' technology varied both by year and region. By testing differences across years and regions, the results suggest that technical efficiency declined in 2007 in both the South and West region during the financial crisis, and has not recovered completely across all cases. Comparatively, the Northeast did not see any changes in mean efficiency as a result of the financial crisis. Notably, the Midwest and South saw the largest decline in mean efficiency from 2002–2007. These results have significant implications in regards to efficiency of spending and shed light on the impact of the financial crisis at the local government level.

Table 2.1: Summary Statistics

Variable	Min	Q1	Median	Mean	Q3	Max
X	3.010×10^{2}	30.490×10^{3}	57.270×10^{3}	1.515×10^{5}	1.208×10^{5}	8.862×10^{6}
Y_1	7.180×10^{2}	3.012×10^{4}	4.648×10^{4}	8.987×10^{4}	7.934×10^{4}	3.849×10^{6}
Y_2	5.200×10^{1}	3.942×10^{3}	7.558×10^{3}	1.752×10^{4}	1.487×10^{4}	8.128×10^{5}
Y_3	1.150×10^{-6}	1.608×10^{-4}	2.345×10^{-4}	2.985×10^{-4}	3.452×10^{-4}	2.068×10^{-2}
Y_4	5.400×10^{-1}	1.199×10^{1}	2.182×10^{1}	4.208×10^{1}	3.951×10^{1}	2.717×10^{3}
Y_5	1.000×10^{0}	5.500×10^{1}	1.900×10^{2}	4.448×10^{2}	4.730×10^{2}	1.162×10^{4}
Y_6	6.480×10^{1}	9.290×10^{1}	9.470×10^{1}	9.421×10^{1}	9.610×10^{1}	9.950×10^{1}
Y_*	1.066×10^{2}	4.884×10^{2}	7.076×10^{2}	1.319×10^{3}	1.177×10^{3}	6.451×10^{4}

Table 2.2: Numbers of Observations With Estimated Hyperbolic Technical Efficiency Equal to 1 in Each Year

		Without Dimension — Reduction —			With Dimension — Reduction —		
Year	n	FDH	VRS	CRS	FDH	VRS	CRS
1997	649	338	43	9	42	7	1
2002	730	355	39	6	35	8	1
2007	746	358	38	9	29	6	1
2012	800	373	36	7	33	6	1

Table 2.3: Results of Convexity Tests (with Dimension Reduction, 1,000 Splits, 1,000 Bootstrap Replications; $p = q = 1$)

	— Input —		— Output —		— Hyperbolic —	
Year	Statistic	p-value	Statistic	p-value	Statistic	p-value
Test #1:						
1997	3.566	0.000	0.653	0.167	1.175	0.003
2002	3.772	0.000	0.683	0.118	1.277	0.005
2007	2.116	0.005	1.141	0.795	3.660	0.000
2012	3.022	0.000	1.222	0.898	4.167	0.000
Test #2:						
1997	0.583	0.000	0.297	0.295	0.529	0.001
2002	0.685	0.000	0.321	0.120	0.469	0.007
2007	0.674	0.000	0.521	0.004	0.751	0.000
2012	0.738	0.000	0.584	0.000	0.817	0.000

Table 2.4: Summary Statistics for FDH Technical Efficiency Estimates (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max	Var
— Input Orientation —							
1997	0.044	0.206	0.331	0.416	0.597	1.000	0.075
2002	0.024	0.174	0.282	0.374	0.505	1.000	0.070
2007	0.036	0.194	0.310	0.379	0.506	1.000	0.062
2012	0.032	0.190	0.287	0.366	0.458	1.000	0.062
— Output Orientation —							
1997	0.182	0.457	0.590	0.609	0.751	1.000	0.041
2002	0.135	0.442	0.567	0.587	0.726	1.000	0.040
2007	0.138	0.421	0.519	0.556	0.687	1.000	0.039
2012	0.119	0.387	0.501	0.540	0.674	1.000	0.042
— Hyperbolic Orientation —							
1997	0.346	0.587	0.705	0.714	0.826	1.000	0.025
2002	0.301	0.575	0.665	0.690	0.811	1.000	0.025
2007	0.259	0.523	0.640	0.660	0.784	1.000	0.028
2012	0.221	0.505	0.623	0.647	0.777	1.000	0.031

NOTE: Statistics for the reciprocals of the output efficiency estimates are given to facilitate comparison with the input-oriented and hyperbolic estimates.

Table 2.5: Tests for Differences in Mean Efficiency across Periods (FDH with Dimension Reduction; $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
1997–2002	-17.955	4.372×10^{-72}	-7.250	4.172×10^{-13}	-13.086	3.980×10^{-39}
2002–2007	-12.014	3.010×10^{-33}	-6.933	4.121×10^{-12}	-11.951	6.393×10^{-33}
2007–2012	4.662	3.136×10^{-06}	7.407	1.288×10^{-13}	10.996	3.991×10^{-28}
1997–2012	-15.482	4.598×10^{-54}	-12.823	1.213×10^{-37}	-18.528	1.235×10^{-76}

Table 2.6: Tests of Differences in Mean Efficiency between Years, by U.S. Census Region (FDH with Dimension Reduction; $p = q = 1$)

	— Input —		**— Output —**		**— Hyperbolic —**	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
— Region 1: Northeast —						
1997–2002	4.947	7.529×10^{-07}	6.795	1.080×10^{-11}	3.586	3.353×10^{-04}
2002–2007	2.597	9.412×10^{-03}	6.685	2.311×10^{-11}	4.859	1.181×10^{-06}
2007–2012	−0.745	4.565×10^{-01}	−5.259	1.446×10^{-07}	−1.253	2.104×10^{-01}
1997–2012	7.464	8.389×10^{-14}	1.160	2.461×10^{-01}	3.926	8.625×10^{-05}
— Region 2: Midwest —						
1997–2002	−0.554	5.795×10^{-01}	−0.962	3.363×10^{-01}	−0.159	8.735×10^{-01}
2002–2007	0.579	5.623×10^{-01}	0.512	6.088×10^{-01}	0.296	7.675×10^{-01}
2007–2012	−3.874	1.071×10^{-04}	−5.038	4.697×10^{-07}	−3.390	6.995×10^{-04}
1997–2012	−6.167	6.976×10^{-10}	−4.385	1.160×10^{-05}	−3.756	1.724×10^{-04}
— Region 3: South —						
1997–2002	−2.672	7.533×10^{-03}	−0.446	6.557×10^{-01}	−0.832	4.055×10^{-01}
2002–2007	0.558	5.766×10^{-01}	0.039	9.688×10^{-01}	0.177	8.592×10^{-01}
2007–2012	3.023	2.499×10^{-03}	0.418	6.761×10^{-01}	2.036	4.175×10^{-02}
1997–2012	−2.746	6.039×10^{-03}	−0.676	4.990×10^{-01}	−1.161	2.455×10^{-01}
— Region 4: West —						
1997–2002	−5.904	3.547×10^{-09}	−2.964	3.033×10^{-03}	−4.026	5.665×10^{-05}
2002–2007	−7.129	1.012×10^{-12}	−4.688	2.760×10^{-06}	−8.842	9.376×10^{-19}
2007–2012	−4.577	4.712×10^{-06}	−6.798	1.058×10^{-11}	−7.289	3.112×10^{-13}
1997–2012	−0.498	6.185×10^{-01}	−3.438	5.861×10^{-04}	−3.305	9.514×10^{-04}

Table 2.7: Test for Separability with Respect to Time (FDH with Dimension Reduction, 1,000 Splits, 1,000 Bootstrap Replications; $p = q = 1$)

	— Input —		— Output —		— Hyperbolic —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
Test #1:						
1997–2002	2.752	0.000	16.912	0.000	7.743	0.000
2002–2007	2.496	0.000	7.298	0.000	4.561	0.000
2007–2012	2.674	0.000	16.328	0.000	7.513	0.000
1997–2012	1.792	0.000	8.451	0.000	4.730	0.000
Test #2:						
1997–2002	0.598	0.000	1.000	0.000	0.983	0.000
2002–2007	0.528	0.000	0.917	0.000	0.748	0.011
2007–2012	0.571	0.000	0.998	0.000	0.820	0.000
1997–2012	0.489	0.000	0.940	0.000	0.739	0.010

Table 2.8: Tests for Separability with Respect to U.S. Census Region (FDH with Dimension Reduction, 1,000 Splits, 1,000 Bootstrap Replications; $p = q = 1$)

	— Input —		— Output —		— Hyperbolic —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
Test #1:						
1997	2.716	0.000	7.805	0.000	4.119	0.000
2002	3.761	0.000	8.457	0.000	4.960	0.000
2007	3.636	0.000	10.790	0.000	5.982	0.000
2012	4.287	0.000	12.169	0.000	6.844	0.000
Test #2:						
1997	0.649	0.000	0.985	0.000	0.905	0.023
2002	0.780	0.000	0.993	0.000	0.956	0.000
2007	0.755	0.000	0.993	0.000	0.947	0.000
2012	0.885	0.000	0.998	0.000	0.984	0.000

Table 2.9: Summary Statistics for Hyperbolic FDH Technical Efficiency Estimates by U.S. Census Region (with Dimension Reduction; $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max	Var
— Region 1: Northeast —							
1997	0.501	0.644	0.735	0.753	0.867	1.000	0.022
2002	0.494	0.646	0.762	0.771	0.906	1.000	0.023
2007	0.479	0.684	0.784	0.792	0.918	1.000	0.023
2012	0.525	0.722	0.817	0.826	0.946	1.000	0.016
— Region 2: Midwest —							
1997	0.545	0.745	0.868	0.856	1.000	1.000	0.017
2002	0.488	0.739	0.841	0.841	0.957	1.000	0.015
2007	0.462	0.779	0.866	0.853	0.952	1.000	0.014
2012	0.406	0.743	0.835	0.824	0.921	1.000	0.016
— Region 3: South —							
1997	0.357	0.679	0.826	0.796	0.917	1.000	0.024
2002	0.328	0.658	0.800	0.779	0.896	1.000	0.024
2007	0.293	0.665	0.804	0.780	0.909	1.000	0.026
2012	0.335	0.660	0.813	0.779	0.901	1.000	0.025
— Region 4: West —							
1997	0.452	0.653	0.756	0.769	0.907	1.000	0.024
2002	0.381	0.624	0.728	0.743	0.880	1.000	0.025
2007	0.391	0.578	0.716	0.718	0.856	1.000	0.029
2012	0.372	0.577	0.725	0.727	0.878	1.000	0.033

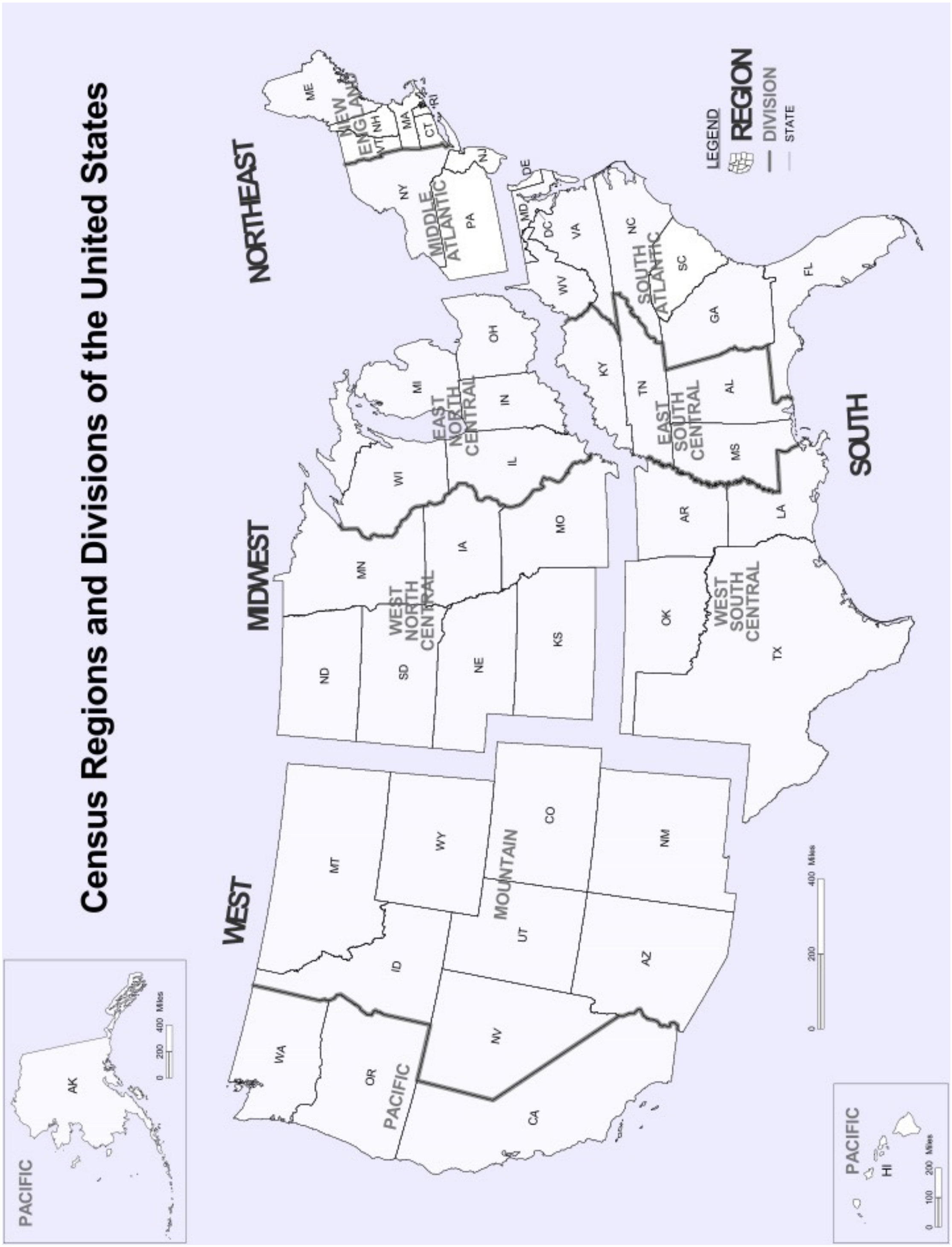
Table 2.10: Summary Statistics for Productivity Estimates by U.S. Census Region (with Dimension Reduction; $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
— Region 1: Northeast —						
1997	3.476	7.031	10.051	12.585	14.396	53.726
2002	2.580	3.913	5.115	7.080	9.987	24.482
2007	1.672	3.193	4.260	7.179	7.599	139.665
2012	2.611	4.926	6.227	7.644	10.286	18.529
— Region 2: Midwest —						
1997	7.039	16.750	24.314	29.843	32.474	381.186
2002	3.097	11.173	13.757	16.226	17.607	222.833
2007	2.641	8.230	10.012	11.791	12.840	120.193
2012	4.087	12.291	15.175	16.702	18.513	154.299
— Region 3: South —						
1997	4.321	13.187	20.198	21.429	27.596	72.887
2002	1.979	8.033	12.528	13.117	15.999	46.587
2007	1.719	5.583	8.979	9.689	12.593	32.663
2012	1.827	8.379	14.012	14.363	18.366	87.970
— Region 4: West —						
1997	5.494	17.880	22.773	26.086	29.540	138.376
2002	1.536	10.752	14.647	17.301	19.287	100.864
2007	1.400	8.031	9.887	11.951	14.027	60.908
2012	2.538	12.562	16.545	18.662	20.973	91.286

Table 2.11: Tests of Differences in Means for Productivity Estimates by U.S. Census Region and Years (with Dimension Reduction; $p = q = 1$)

	— **Productivity** —	
Period	Statistic	p-value
	— Region 1: Northeast —	
1997–2002	-6.984	2.862×10^{-12}
2002–2007	0.053	9.574×10^{-01}
2007–2012	0.254	7.991×10^{-01}
1997–2012	-5.712	1.115×10^{-08}
	— Region 2: Midwest —	
1997–2002	-3.729	1.924×10^{-04}
2002–2007	-4.441	8.954×10^{-06}
2007–2012	6.517	7.167×10^{-11}
1997–2012	-3.763	1.678×10^{-04}
	— Region 3: South —	
1997–2002	-13.408	5.421×10^{-41}
2002–2007	-12.761	2.706×10^{-37}
2007–2012	15.093	1.795×10^{-51}
1997–2012	-10.132	4.001×10^{-24}
	— Region 4: West —	
1997–2002	-12.276	1.222×10^{-34}
2002–2007	-7.851	4.130×10^{-15}
2007–2012	14.918	2.523×10^{-50}
1997–2012	-6.737	1.611×10^{-11}

Figure 2.1: Census Regions of the United States



Chapter 3

The Effect of Proposition G on Housing Quality and Tenant Filtering in San Francisco

3.1 Introduction

Finding rental housing in San Francisco can be a challenge. Due in part to supply constraints and increased demand for rental housing, renters face a limited stock of affordable units. According to Rent Jungle, as of January 2018 the average rent for a one bedroom apartment in San Francisco was over \$3,400 a month and nearly \$4,500 for a two bedroom unit. The city has attempted to deal with the issue of housing adequacy and affordability since the late 1970s when it initially implemented rent control policies. Since San Francisco does not track the housing stock over time, the total number of controlled units is unknown. In 2014, urban research group Spur estimated there are approximately 172,000 controlled units in San Francisco. This constitutes nearly 72 percent of the total stock of rental housing and would suggest that rent control coverage is extensive.

With citywide rising rents, landlords of rent-controlled buildings face increasing incentives to evade the rent control ordinance. In the 1990s these efforts led to frequent exploitation of loopholes in the laws, often by way of owner move-in (OMI) evictions and subsequent condo conversions of units in controlled buildings via tenancies-in-common. The city has attempted to deal with speculative behavior on the part of landlords and other third parties by passing regulations, notably Proposition G. San Francisco passed Proposition G in 1998 in an effort to stymie the exploitation of these loopholes by increasing the costs associated with the use of these no-fault evictions to landlords and also restricting their usage in controlled buildings. The goal of Proposition G was to prevent further reduction in the controlled stock and protect tenants currently residing in these buildings.

This paper analyzes the effect of Proposition G on rental housing quality in San Francisco. While the regulation was passed in an effort to protect tenants residing in controlled buildings and to preserve the existing stock of affordable housing, it may have resulted in unintended effects. Following the regulation, landlords faced higher opportunity costs of providing rent-controlled housing and greater restrictions on their ability to exit the controlled market. I expect this policy to have the effect of decreasing housing quality as landlord behavior is constrained, resulting in lower building and maintenance expenditures. Additionally, I expect that increased tenant protection from eviction will slow the rate of downward filtering which occurs as a result of the aging of the stock of housing and rental property turnover.

Theory suggests that rent control policies are likely to result in one or both of the following: lower rental unit quality or a change in the contractual agreement between landlords and tenants. While the empirical literature is mixed, there is an obvious disincentive for a landlord to maintain the unit if a binding price ceiling exists.

To get around these caps, landlords may have prospective renters bid ("key money") on the unit in an attempt to recoup the losses suffered from the rent control mandates. Cheung (1975) looks at the unintended consequences following the 1921 Rent Control Ordinance in Hong Kong. While the legislative intent of the bill was to control the existing rents and to encourage construction on unoccupied land, he notes that landlords will evict their tenants so long as the expected profits from reconstruction exceed the rental stream that can be obtained from the prospective tenant. The empirical evidence suggests that this ordinance encouraged reconstruction as a means to escape the rent control mandates.

Cheung (1979) examines the effects of rent control on housing reconstruction in Hong Kong. When landlords were only permitted to evict tenants for the purpose of reconstruction, this had an unintended effect of causing frequent demolition of both old and new buildings alike. Through demolition and reconstruction, they were able to circumvent the rent control ordinance in Hong Kong. These findings highlight the issue with rent control regulation and the unintended effects it can have. This behavior leads to considerable waste and is an inefficient use of resources. Landlords of controlled buildings in San Francisco and other markets with rent control likely face similar incentives in regards to demolition and reconstruction. Proposition G aimed to make this practice more difficult in San Francisco by imposing higher costs on this behavior. In addition to its unintended effects, rent control also may impact the allocation of resources. Olsen (1972) finds that occupants of New York City rent-controlled housing consume around four percent less housing services than they would have in the absence of rent controls. He also finds that poorer families received larger benefits than richer ones. While the issue of whether rent control increases homelessness is highly debated in the literature, Early and Olsen (1998) cannot reject the hypothesis that rent control has no net effect on homelessness.

Additionally, theory suggests that units in controlled buildings would be of lower average quality compared to similar units in uncontrolled buildings. This can be attributed primarily to the differing incentives faced by landlords of these respective property types. Housing quality is difficult to observe in the data, but since the intent of rent control is to ensure adequate and affordable housing, quality is especially relevant in this context. Mengle (1985) examines whether rent control leads to lower quality than would have otherwise been observed. He tests whether rent control causes landlords to reduce their maintenance expenditures and whether the effects of rent control become more pronounced over time. The results from his empirical analysis suggest that rent control leads to lower quality units and the effect tends to worsen with age. While he uses a national sample of housing units to examine the impact of rent control on quality, it is likely that this result also holds for smaller markets.

Similarly, Moon and Stotsky (1993) examine the impact of rent control on housing quality using longitudinal data from New York City. Using a nonstationary heterogenous Markov model and a mover-stayer structure, their results provide evidence that rent control does lead to a deterioration in housing quality. However, they note that additional research is needed on this issue. Gyourko and Linneman (1990) examine rent control and its impact on housing quality across different building sizes in New York City. They estimate a logit model and control for observable differences in buildings including rent control status, building age, and building type (e.g. high-rise). Their results suggest that rent control had a larger negative effect on quality for relatively smaller buildings. Unsurprisingly, the effect is largest for buildings built prior to 1947.

Rent control also has the effect of slowing filtering in a city. Filtering is the idea that as houses age, rents must decrease in order to get people to live in them. With rent control, landlords who are renting below market rates have no incentive to lower the price further, even with age and wear and tear of the unit. Mills and Hamilton (1994) note that filtering is a consequence of income growth and deterioration of housing units over time. With rent control, it can be difficult to observe filtering as tenants may remain at their controlled units longer than they would have otherwise. Munch and Svarer (2002) show that tenancy mobility is severely reduced in the Danish rent control market. Results from Diamond et al. (2019) show that expansion of rent control policies in San Francisco led to a reduction in the supply of housing on the behalf of landlords, thereby raising prices in the uncontrolled sector. Their findings also suggest that rent control leads to longer tenancy duration and do not necessarily benefit those that the rent control policies are intended for.

If downward filtering holds, wealthier households will move into the newest housing stock and the remaining housing filters down until households with the least means move into the second-lowest quality housing and the lowest quality housing drops out of the stock. It is plausible that San Francisco's housing and rent control policies have the effect of restricting tenant mobility and therefore impede downward filtering. Rent control advocates often argue that rent control leads to a mixing of both wealth and races, while economic theory does not necessarily support this argument for reduced segregation. Glaeser (2002) finds that rent control in wealthier areas has enabled poor renters to live in these cities, while declining cities in New Jersey with rent control have increased isolation of the poor. He argues for supply-side policies as a solution to reduce segregation as opposed to rent control mandates.

Using household-level longitudinal data from the American Housing Survey (AHS), I utilize a stochastic rent frontier to estimate the effect of unit and location-specific attributes on annual contract rent, controlling for differences in rent control coverage. I assume that housing quality is independently distributed and a linear function of observables, following Battese and Coelli (1995). Observables include whether the building owner lives in the building, whether the building is rent-controlled, tenant and household characteristics, and maintenance indicators. To estimate the causal impact of the regulation change, I include interactions between the post-regulation period, the presence of a live-in building owner, and the treatment group, San Francisco rent-controlled buildings. Assuming landlords are profit maximizers and are obtaining the maximum possible rent (subject to constraints), my results suggest that following Proposition G, owner-occupied rent-controlled buildings in San Francisco saw lower mean quality, reflecting higher post-regulation costs. Furthermore, my findings provide evidence that rent control slows downward filtering, consistent with findings in Diamond et al. (2019).

The paper is organized as follows. In Section 3.2, I provide details on rent control in San Francisco and describe the motivation behind Proposition G and what the regulation entails. In Section 3.3, I detail the statistical model and discuss its advantages over the traditional hedonic approach. I detail the data in Section 3.4 and present the findings and results in Section 3.5. Lastly, in Section 3.6, I conclude and summarize the main findings of the paper and discuss the relevant policy implications.

3.2 Rent Control Background

Rent control arrived in San Francisco with the Rent Stabilization Ordinance in 1979 following pressure over rising housing costs from city residents. The goal of the rent stabilization ordinance was to ensure a stock of "adequate and affordable housing for low to middle-income households." Although adequacy is not explicitly defined, the intent of rent control was to provide lower-cost housing options to residents in San Francisco. One of the key provisions of the Rent Stabilization Ordinance is that rent control only applies to buildings built prior to June 13, 1979. There are a few exceptions to this, and single-family homes as well as condos face limited rent control coverage in most cases.[1] Furthermore, the city expanded rent control in 1994 with the goal of increasing the stock of controlled units to include multiunit buildings with four or less units ("mom and pop landlords"), which were previously exempt from the ordinance.

Landlords of controlled buildings are permitted to increase tenants' rents annually for those that choose to renew, but these increases are subject to a cap determined by the Rent Board and tied to the rate of inflation for the Bay Area. With new tenants, landlords are permitted to set a new base rent without restrictions, although tenants in compliance may not be denied a lease renewal. According to the San Francisco Tenants' Union, landlords that perform maintenance or capital improvements may pass on a portion of those costs to the tenants, but this amount is capped and tenants are permitted to petition for exemptions if this occurs. It should be noted that while normal annual increases in contract rent are capped, landlords are permitted to increase rent following the eviction of a tenant at-fault. Once a tenant leaves and is proven to be at-fault, the landlord is able to raise the rent to the market rate, although the unit remains under rent control for subsequent increases.

The San Francisco Rent Board manages and oversees evictions in the city. Evictions of tenants in both uncontrolled or controlled buildings require just-cause on behalf of the landlord and are divided into two types: at-fault and no-fault. There are 16 grounds for eviction in total and all eviction notices must be given to and approved by the Rent Board before occurring. Proposition G (1998) imposed restrictions on the use of OMI evictions, a type of no-fault eviction. An OMI eviction occurs

[1] California passed the Costa Hawkins law in 1995 which set requirements for cities with rent control. The law protects a landlord's right to raise rents after a tenant moves and exempts units constructed after February 1995 as well as single-family homes and condos from rent control.

when the landlord or a close relative intends to move into the unit. Along with the complying with the grounds for evictions, landlords utilizing these no-fault evictions are also required to compensate tenants, as is the case with all no-fault evictions. This amounts to roughly $5,000 per tenant, but can be up to $15,000. Furthermore, as of 2015, following an OMI eviction, there is a five year tenancy decontrol period in which the evicted tenant can re-rent the unit at the same rate prior to eviction (subject to allowable increases).

Prior to the passing of Proposition G, OMI evictions were popular with landlords for several reasons. Under this eviction type, the owner could evict a tenant under what effectively amounts to "good faith." This means that the Rent Board took the landlord at their word that they or a close relative would eventually move-in and occupy the unit. It is likely that this did not always happen and in some cases, it is plausible that a landlord would let their building sit empty or permit unrelated parties to occupy vacant units. Furthermore, OMI evictions enabled landlords to exploit shared ownership of buildings via tenancies-in-common. A tenancy-in-common has multiple owners of a single building and were exempt from the condo conversion laws.

As Capps (2014) notes, the number of OMI evictions peaked in 1998, the year that Proposition G was passed in San Francisco. Table 3.1 lists the total OMI eviction filings by fiscal year since the San Francisco Rent Board began tracking them. The total number of OMI filings increased up until the regulation was passed in 1998 and then steadily declined in the years following. Prior to this binding constraint, landlords commonly utilized OMI evictions to change the building into a tenancy-in-common, and then proceeded towards condo conversion, contingent upon receiving a city granted demolition permit. Proposition G most notably caps the number of OMI evictions by landlords to one per building, as well as increases the occupancy and move-in requirements making it more difficult for landlords to exploit loopholes in the law. The key amendments to the use of OMI evictions, appearing in Proposition G, follow.

3.2.1 Proposition G (1998) Amendments

The intent of this municipal regulation was to prevent landlords from exploiting loopholes in the conversion and reconstruction law. The key amendments to the usage and implementation of OMI evictions for the city of San Francisco is as follows. The full referendum can be found in Appendix (A).

(i) Landlords may only use one OMI eviction per building.

(ii) If a landlord intends to use an OMI eviction for a relative move-in, they must already reside in the building.

(iii) If there is already a vacant and comparable unit in the building, the use of an OMI eviction is limited.

(iv) Once an OMI eviction is used, the landlord or relative must move in within 3 months and continuously occupy the unit for 3 years.

(v) If the landlord uses an OMI eviction, they must offer the evicted tenant another rental unit at a comparable rate.

(vi) Use of OMI evictions on senior, disabled, or terminally ill tenants is not permitted.

This regulation adds to the ever-growing number of housing regulations in the San Francisco market. By restricting the number of OMI evictions to one per building, this regulation prevents landlords from exploiting the use of this no-fault eviction. Furthermore, by requiring the landlord to occupy the unit continuously for three years, this regulation forces landlords to use OMI evictions for non-speculative behavior. Proposition G also increases protections for evictions of the protected classes.[2]

3.3 Statistical Model

Stochastic frontier analysis (SFA) was simultaneously introduced by Aigner et al. (1977) and Meeusen and van den Broeck (1977) and is frequently used in many different applications including health, finance, and agriculture. It has also been utilized to model the production of rental housing services. Caudill (1993) uses SFA to examine how much rents would adjust in the absence of rent control in New York City. He asserts that the use of SFA is preferable to the traditional hedonic approach for several reasons, primarily being that the hedonic approach implicitly assumes that rents in the uncontrolled sector would not change upon decontrol. This is a strong assumption and may not be realistic. He finds that rents would substantially adjust if rent control were to be abolished, rising between 22 and 26 percent in the controlled sector and falling between 22 and 25 percent in the uncontrolled sector.

Following the literature, I use SFA to model the production of rental housing services and estimate the impact of Proposition G on housing quality. For a rental unit with given observable characteristics, I define output as the maximum attainable annual contract rent. The inputs to annual contract rent include unit-specific attributes, such as the number of rooms and rent control status, as well as attributes of the surrounding area. Conditional on rent control status, I expect annual contract rent to reflect the full consumption value of the unit. Let R_{it} represent the annual

[2]Buildings with OMI evictions face additional decontrol regulations; see the San Francisco Rent Board for details.

contract rent of unit i at time t

$$R_{it} = f(X_{it}, Z_{it}; \beta_1, \beta_2) exp(\varepsilon_{it}), \tag{3.1}$$

where X_{it} is a vector of observable unit characteristics, Z_{it} is a vector of location-specific attributes including neighborhood effects, and ε_{it} is a discount factor. Specifying a log-linear production function, (3.1) can be rewritten as

$$\begin{aligned} \log R_{it} &= \mathbf{X_{it}}\beta_1 + \mathbf{Z_{it}}\beta_2 + \varepsilon_{it} \\ \varepsilon_{it} &= v_{it} - u_{it} \end{aligned} \tag{3.2}$$

where v_{it} are identically and independently distributed normal random errors and u_{it} are non-negative independently distributed random errors.

I estimate a single frontier, controlling for differences in rent control coverage. My identification consists of two key assumptions. First, I assume that landlords are profit maximizers and are therefore charging the maximum possible rent in either the controlled or uncontrolled sector. Hence, any differences in rents between two otherwise identical units, in either the uncontrolled or controlled sector, can be attributed to differences in observable rental housing quality. When $u_{it} = 0$, (3.2) defines the maximum rent frontier and any differences between rental units may be attributed to random noise. For units that rent at a relative discount $u_{it} > 0$, thereby capturing disquality. I also assume that the AHS data are representative of the true housing stock. Observable housing quality is estimated as

$$q_{it} = exp(-u_{it}) \tag{3.3}$$

which is based on its conditional expectation following Battese (1993) and Battese and Coelli (1995).

From (3.3), let quality be modeled as a linear function of observables

$$\begin{aligned}q_{it} = {} & A_{it}\delta_1 + \delta_2\text{Post}_{it} + \delta_3\text{Treatment}_{it} + \delta_4\text{Owner}_{it} + \\ & \delta_5(\text{Treatment}_{it} \times \text{Post}_{it}) + \delta_6(\text{Owner}_{it} \times \text{Post}_{it}) + \\ & \delta_7(\text{Treatment}_{it} \times \text{Owner}_{it}) + \delta_8(\text{Treatment}_{it} \times \text{Post}_{it} \times \text{Owner}_{it}) + \epsilon_{it}\end{aligned} \tag{3.4}$$

where A_{it} is a vector of controls including unit and tenant characteristics, *Post* is an indicator for the post-regulation period, *Treatment* is an indicator for rent-controlled buildings in San Francisco city, *Owner* is an indicator for owner-occupied buildings, and ϵ_{it} is a random error. The marginal effect of *Treatment* on q_{it} from (3.4) is given by

$$\frac{\partial q_{it}}{\partial \, Treatment} = \begin{cases} \delta_3 & \text{if } Post = 0 \;\&\; Owner = 0, \\ \delta_3 + \delta_5 & \text{if } Post = 1 \;\&\; Owner = 0, \\ \delta_3 + \delta_7 & \text{if } Post = 0 \;\&\; Owner = 1, \\ \delta_3 + \delta_5 + \delta_7 + \delta_8 & \text{if } Post = 1 \;\&\; Owner = 1. \end{cases}$$

Following Proposition G, I expect controlled buildings in San Francisco to be associated with lower quality, reflecting lower building maintenance and upkeep. If downward filtering holds, as buildings age quality should be lower. Additionally, older tenants and those with college education should be associated with higher quality, whereas minority and low-income tenants should reside in units with lower associated quality.

Before estimating a stochastic rent frontier, it is necessary to identify the presence of u_{it}, or technical inefficiency. Testing for the presence of u_{it} is necessary because if $u_{it} = 0$, then $\varepsilon_{it} = v_{it}$ in (3.2), which indicates a symmetric error term, suggesting that the use of stochastic frontier is not be appropriate. To do this, a test can be performed on the residuals from a simple ordinary least squares (OLS) estimation, proposed by Schmidt and Lin (1984). Using the second and third sample moment, m_2 and m_3, respectively, the test statistic for the presence of technical inefficiency, can be written as

$$b = \frac{m_3}{(6m_2^3)/n}^{\frac{1}{2}} \sim N(0, 1). \tag{3.5}$$

If m_3 is less than zero, this may indicate the presence of technical inefficiency. If m_3 is greater than zero, it may be an indicator of model misspecification and further examination is necessary. The residuals show a negatively skewed distribution. This suggests the presence of u_{it} and the use of a stochastic frontier model is justified.

3.4 Data

Data used for estimation are obtained from the American Housing Survey (AHS) metropolitan surveys. The AHS asks households about their housing accommodations, the surrounding area, as well as information regarding their household and household composition. The data are longitudinal with households being surveyed once to several times over the life of the survey. These data are presumed to be representative of the existing stock with units being added and removed to reflect changes over time. I utilize data from the San Francisco metropolitan statistical area (MSA) for the following years: 1985, 1989, 1993, 1998, and 2011. These data allow me to identify the municipality in which the rental units are located, including those in San Francisco city. I include only renter-occupied units and omit any observations that are government projects or that receive government subsidies. A full list of variables and variable descriptions can be found in Table 3.2.

3.4.1 Hedonic Component

In order to estimate the hedonic pricing component, it is necessary to control for any observables that may impact annual contract rent. These include characteristics of the rental unit such as building age, number of rooms, and number of bathrooms. Increases in the number of rooms or bathrooms is likely positively correlated with unit size and hence, should result in higher rent. I account for any possible nonlinear effects by including quadratic terms for building age and number of rooms. Additionally, I control for the presence of central heat and central air conditioning. It is plausible that households value these attributes even with the temperate San Francisco climate. I also account for the presence of a balcony, working fireplace, and the existence of covered parking, all of which I expect to have a positive impact on annual contract rent.

To control for housing an neighborhood satisfaction, I include two self-reported measures. The AHS question asks households "On a scale from 1–10, how satisfied are you with the housing unit?" and similarly for the neighborhood survey question. Following the AHS question regarding rent control coverage, I include an indicator for rent-controlled units.[3] I also distinguish between building sizes. Since "mom and pop" landlords were subject to rent control mandates in 1994, I control for the presence of large buildings (e.g. buildings with more than four units).

In addition to observable unit characteristics, I control for neighborhood effects and other characteristics of the surrounding area. To capture differences between suburban and urban areas, I include the zone-level mean number of rooms. I expect

[3] AHS survey participants self-report whether there unit is rent-controlled.

increases in the mean number of rooms to have a negative impact on contract rent since this reflects lower land prices and likely a more suburban area. Moreover, I control for the zone-level mean household size, mean household income, and mean building age. Larger households are more likely to reside in relatively suburban areas and hence, increases in this measure should be associated with lower rent. For mean household income, the effect may go either way if wealthier households choose to locate in the city center, closer to amenities, or move outside of the city to more suburban areas. To account for differences in rent control coverage, I use the zone-level mean building age.

3.4.2 Quality Component

To examine filtering in the quality component estimation, I control for observable unit and tenant characteristics, maintenance indicators, as well as post-regulation period and treatment group indicators and their interactions to account for the impact of Proposition G on housing quality. I include a rich set of attributes concerning the head of household including whether they are male, older (e.g. 65 years or older), younger (e.g. less than 30 years), married, college-educated, and black. If downward filtering holds in San Francisco, the newest housing filters down from the wealthiest to the poorest households, until the lowest quality household falls out of the housing stock. I expect that older households will tend to be associated with higher housing quality and the opposite to be true for units with a younger head of household. It is also plausible that married and college-educated tenants be associated with higher housing quality, possibly due to having greater financial resources.

To control for other factors that may impact housing quality, I use several different maintenance and adequacy indicators. These include whether there are leaks inside the unit, whether there is broken plaster, the presence of visible cracks, and whether there has been a sewage breakdown in the last 90 days. These should all be associated with lower quality. Additionally, the AHS provides a composite adequacy rating which I also include. To account for differences in maintenance incentives, I control for the presence of the owner in the building. Owners that live in their controlled building likely have greater incentive to maintain it. I also include indicators for the post-regulation period and the treatment group, San Francisco controlled buildings.

Table 3.3 presents selected descriptive statistics for the sample. Notably, there is nearly a 100 year old range between the oldest and newest building in the sample, with a minimum and maximum building age of 1919 and 2010, respectively. Nearly 40 percent of rental units reside in controlled buildings. It should be noted that this includes controlled buildings in both San Francisco city as well as the surrounding areas. Additionally, there is considerable variation in annual contract rent, likely reflecting differences between controlled and uncontrolled buildings, with a minimum annual contract rent of $2,692, compared to a maximum of $34,684. The median and mean contract rent are $11,937 and $13,041, respectively, reflecting a right-skewed distribution.

In terms of observable unit characteristics, the median and mean number of rooms are four and 3.83, respectively. This suggests the sample consists primarily of smaller units, which can be expected since the San Francisco area is primarily urban. Consistent with the number of rooms, the mean number of bedrooms and bathrooms are around two and one, respectively. Approximately 77 percent of units reside in large buildings, or buildings with more than four units. The average building age in the sample is a little over 45 years old at the time of sample. Finally, 44 percent of buildings have an on-site building owner.[4]

Furthermore, there is evidence of maintenance issues in the sample. Nine percent of rental units display visible cracks, 16 percent of units have inside leaks of some form, and seven percent have walls with broken plaster. While there are obvious maintenance issues, median housing and neighborhood satisfaction ratings are relatively high, averaging an eight out of ten. Only ten percent of the sample consists of older head of households, whereas 22 percent of the units have a younger head of household. The majority of tenants are older than 30 and less than 65. Nine percent of households have a black head of household, compared with 54 percent with a male head of household. Additionally, the mean household size is around two.

3.5 Empirical Results

Results from estimation of the hedonic pricing model component are presented in Table 3.4. In column 1, I present the results from the baseline estimation, without including any interaction effects in the quality component estimation. The majority of coefficients have the expected signs. The coefficient on building age is negative and significant at one percent. Similarly, the coefficient on the quadratic building age term is positive and significant at five percent, suggesting a nonlinear effect of building age on annual contract rent. Increases in the number of rooms and bathrooms both positively impact contract rent, consistent with initial expectations. Units with covered parking rent for four percent more. Unsurprisingly, units with a working fireplace and balcony rent for approximately nine and seven percent more respectively.

In addition to the observable unit characteristics, I control for neighborhood effects in column 1. Both the mean number of rooms and mean household size are

[4]Unfortunately, I am not able to observe whether the owner has utilized an owner move-in eviction in their building.

negative and statistically significant at one percent, consistent with the expected sign. A fifty percent increase in the zone-level number of rooms (e.g. going from two to three bedrooms) results in nearly a 28 percent increase in the response variable. This likely reflects the scarcity of land and the housing supply constraints in the area. It may also indicate that the area is becoming more suburban. The sign on the coefficient of mean household income is positive. This indicates that increases in income are associated with higher rents, suggesting a positive income elasticity. The coefficient on the mean building age is positive, but insignificant.

The results from the estimation of the quality component from the baseline model are presented in column 1 of Table 3.5. This model specification does not include any interactions capturing the regulation change. The coefficient on the large building term is negative, suggesting that lower quality is associated with buildings with more than four units. This may be due to higher maintenance costs or other difficulties with maintaining a larger space. The coefficients on older and younger are the expected signs and statistically significant at one percent. Older tenants are associated with higher quality units which may due to greater disposable income. By contrast, the coefficient on the younger term is negative suggesting that tenants younger than 30 years are associated with overall lower housing quality.

Results from the full model estimation of the hedonic pricing model component are presented in column 2 of Table 3.4. Most of the signs are as expected and similar to those in the baseline estimation. The coefficient on rent control is negative and statistically significant at ten percent, indicating that rent-controlled units have lower annual contract rent compared to uncontrolled units. This is to be expected. Moreover, increase in both house and neighborhood satisfaction measures are associated with higher annual rent, which is consistent with initial expectations. The signs and magnitude of the coefficients for each of the neighborhood effects terms are also similar to the baseline model in column 1.

In column 2 of Table 3.5 I present the results from the full model estimation of the quality component including indicators for the post-regulation period, treatment group, owner-occupied buildings, and their interactions. The coefficient on the *Treatment* indicator provides a quality relation for controlled units in San Francisco without a live-in building owner and prior to Proposition G and is negative and statistically significant at one percent. Notably, the coefficient on the rent-controlled indicator variable is positive, suggesting rent controlled units in the San Francisco MSA area are associated with higher overall housing quality. Since this indicator only identifies whether a unit is rent-controlled, this is likely picking up higher quality rent-controlled units outside of the San Francisco city.

The marginal effect of *Treatment* on housing quality in the post-regulation period without an on-site owner is the sum of the coefficients on *Treatment* and *Treatment* $\times$ *Post*, around -0.17. This result suggests that following Proposition G, rent-controlled buildings in San Francisco are associated with lower housing quality, relative to the surrounding MSA, prior to Proposition G. This is consistent with initial expectations and may provide evidence reflecting changes in landlord maintenance and upkeep behavior. The main effect concerns the impact of San Francisco controlled buildings with a live-in building owner, following Proposition G. Controlling for differences in whether the building owner lives on-site is important since it may explain important differences in quality. Moreover, several stipulations in the Proposition G regulation specifically address the presence of an on-site owner. Compared with the marginal effect of *Treatment* prior to regulation with an on-site owner, quality is also lower, around -0.1245. This is significant at ten percent and suggests that even rent-controlled buildings with on-site owners are associated with lower housing quality prior to the regulation change.

In addition to examining the impact of the regulation change on housing quality, I also account for differences in whether the building owner lives on-site. The marginal effect of *Treatment* on housing quality in the post-regulation period, with an on-site owner, is the sum of the coefficients on *Treatment*, *Treatment* $\times$ *Post*, *Treatment* $\times$ *Owner*, and *Treatment* $\times$ *Post* $\times$ *Owner*. This is approximately -0.28, suggesting that following Proposition G, housing quality for the treatment group is lower when the building owner lives on-site. It may be the case that the building owner no longer believes that the property is worth maintaining for potential resale. In addition to these findings, I present the distribution of housing quality in Figure 3.2. Similar to Galbraith (1998), the distribution shows a strong left skew.

3.6 Discussion & Conclusion

This paper examines the effects of Proposition G on housing quality and tenant filtering in San Francisco. San Francisco and surrounding cities initially implemented rent control policies in an attempt to quell concerns from local residents over rising rents. While San Francisco rent control policies were intended to ensure a stock of adequate and affordable rental housing, the actual efficiency gains are mixed, with Early (2000) finding that tenants in controlled units would be better off if the controls had never been established. Landlords would also be better off and have attempted to circumvent the ordinance in several ways, one of which being through OMI evictions. With rising OMI evictions and frequent use of condo conversions, San Francisco passed Proposition G in an attempt to prevent landlords from abusing this no-fault eviction and therefore maintain the stock of rental housing.

The results suggest that the impact of Proposition G on San Francisco housing quality is negative and significant. In the hedonic component, attributes such as a balcony and working fireplace positively impact contract rent, consistent with theory. Moreover, my findings for the quality component estimation confirm the filtering hypothesis. Older tenants are associated with higher quality housing, while units with a younger head of household occupy lower quality units. Following the regulation, the results suggest that owner-occupied controlled buildings in San Francisco saw lower quality.

Findings from Moorhouse (1972) suggest that in cases of rent control where an entrepreneur anticipates a lower net revenue stream, it is optimal to adjust maintenance to a lower level. He also notes that it is optimal to further adjust maintenance levels given any unanticipated shocks, continuing until the controlled structure is decontrolled. These results are consistent with my findings. Furthermore, Arnott and Igarashi (2000) find that while most models assumption competitive rent control markets, this is overly simplistic. Using a monopolistically competitive model, he notes that while mild rent control is beneficial, severe rent control policies do more harm in terms of increasing inefficiencies regarding matching. These may have the impact of restricting or slowing downward filtering in the market.

Proposition G was passed in 1998 in an attempt to stop landlords from exploiting OMI evictions in an attempt to condo conversions in the city. Recent literature has shown that rent control leads to tenants staying for longer durations as well as causing rents to increase in the uncontrolled sector. The results from this paper support the idea that rent control slows downward filtering and ultimately leads to lower quality over time. These results highlight the need for alternative housing policies that support the goal of ensuring adequate and affordable housing, without the unintended effects.

Table 3.1: OMI Eviction Notices

Fiscal Year	Notices
87 - 88	522
88 - 89	564
89 - 90	545
90 - 91	469
91 - 92	356
92 - 93	293
93 - 94	344
94 - 95	361
95 - 96	481
96 - 97	1,074
97 - 98	1,410
98 - 99	1,200
99 - 00	937
00 - 01	991
01 - 02	594
02 - 03	422
03 - 04	364
04 - 05	288
05 - 06	248
06 - 07	210
07 - 08	159
08 - 09	143
09 - 10	127
10 - 11	139

Table 3.2: Description of Variables

Variable Name	Description
Dependent Variable	
rent	Annual contract rent
Unit Characteristics	
age	Age of unit at time of survey
numrooms	Number of rooms
numbaths	Number of full bathrooms
numbeds	Number of bedrooms
large	Indicator if multiunit structure with more than four units
central heat	Indicator if central heat
central ac	Indicator if central ac
balcony	Indicator if balcony/porch
fireplace	Indicator if working fireplace
parking	Indicator if parking included in rent
rentcontrol	Indicator if unit is rent-controlled
ownerhere	Indicator if the building owner resides in the building
housesatisfaction	Self-reported measure of housing satisfaction (1-10 rating)
neighbsatisfaction	Self-reported measure of neighborhood satisfaction (1-10 rating)
Household Characteristics	
old	Indicator if head of household 65 years or older
young	Indicator if head of household 30 years or younger
black	Black head of household
numper	Number of persons in the household
male	Male head of household
hhincome	Total household income
married	Married head of household
college	College-educated head of household
Quality Components	
insideleaks	Indicator if leaks from inside home
brokenplaster	Indicator if broken plaster inside unit
visiblecracks	Indicator if visible cracks inside unit
sewbreakdown	Indicator if sewage breakdown in last 90 days
adequate	Housing adequacy index (ranging from adequate to severely inadequate)
Neighborhood Effects	
avgrooms	Zone-level average number of rooms
avgper	Zone-level average household size
avgincome	Zone-level average household income size
avgage	Zone-level average structure age
Geography	
MSA	MSA the unit is located in
centralcity	Indicator if unit located in the city of San Francisco
county	County the unit is located in
zone	Economically homogeneous area with 100,000 or more in population

Table 3.3: Selected Descriptive Statistics

Variable Name	Min	Q1	Median	Mean	Q3	Max
Unit Characteristics						
Annual contract rent ($)	2692	9294	11937	13041	15403	34684
Built year	1919	1930	1955	1949	1965	2010
Building age	1	24	44	45.250	66	92
Number rooms	2	3	4	3.826	4	15
Number bedrooms	1	1	1	1.572	2	10
Number bathrooms	1	1	1	1.096	1	3
Large building	0	1	1	0.768	1	1
Rent-controlled	0	0	0	0.360	1	1
Owner in building	0	0	0	0.439	1	1
House satisfaction	1	6	8	7.480	9	10
Neighborhood satisfaction	1	6	8	7.420	9	10
Maintenance Indicators						
Visible cracks	0	0	0	0.085	0	1
Inside leaks	0	0	0	0.160	0	1
Broken plaster	0	0	0	0.065	0	1
Adequacy index	1	1	1	1.121	1	3
Neighborhood Effects						
Mean number rooms	3.244	4.349	5.062	4.831	5.426	6.597
Mean household size	1.710	2.070	2.440	2.362	2.648	3.441
Mean building age	18.050	32.950	46.850	46.350	57.750	74.030
Mean log household income ($)	10.570	10.990	11.240	11.220	11.430	11.940
Tenant Attributes						
Log household income ($)	0.250	10.250	10.758	10.699	11.269	14.216
Older head of household	0	0	0	0.106	0	1
Younger head of household	0	0	0	0.320	1	1
Black head of household	0	0	0	0.092	0	1
Male head of household	0	0	1	0.541	1	1
Household size	1	1	2	2.118	3	14

* All dollar values are in 2010 US dollars.

Table 3.4: Selected Results from the Hedonic Component Estimation

Dependent Variable: Log annual contract rent	— **Renter Sample** — (1) Baseline	(2) With interactions
Unit Characteristics		
Building age	−0.0967***	−0.0978***
	(0.0221)	(0.0219)
Building age^2	0.0104**	0.0106**
	(0.0043)	(0.0043)
Number rooms	0.3826***	0.3904***
	(0.0986)	(0.0993)
Number rooms^2	−0.0369	−0.0399
	(0.0370)	(0.0373)
Number bathrooms	0.1511***	0.1505***
	(0.0171)	(0.0171)
Rent-controlled	−0.0352**	−0.0267*
	(0.0150)	(0.0151)
Large multiunit	0.0038	0.0015
	(0.0127)	(0.0127)
Covered parking	0.0400***	0.0401***
	(0.0085)	(0.0085)
Working fireplace	0.0908***	0.0912***
	(0.0101)	(0.0101)
Balcony (porch)	0.0689***	0.0691***
	(0.0082)	(0.0082)
Central heating	0.0179**	0.0176**
	(0.0078)	(0.0077)
House satisfaction	0.0294**	0.0298**
	(0.0122)	(0.0122)
Neighborhood satisfaction	0.0344***	0.0327***
	(0.0096)	(0.0096)
Neighborhood Effects		
Mean number rooms	−0.5526***	−0.5468***
	(0.0847)	(0.0847)
Mean household size	−0.2334***	−0.2382***
	(0.0578)	(0.0584)
Mean household income	0.4080***	0.4075***
	(0.0427)	(0.0425)
Mean building age	0.0217	0.0178
	(0.0303)	(0.0301)
Location fixed effects	Yes	Yes
Year fixed effects	Yes	Yes

This table presents selected results from estimation of the stochastic frontier hedonic component in (3.1). Column 2 presents results for the hedonic component, controlling for interaction effects in the quality component specification. The sample consists of 5,097 observations from 1985 to 2011 and includes 5 counties in total. All model specifications include year and location fixed effects. Standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 3.5: Selected Results from the Quality Component Estimation

Dependent Variable:	— **Renter Sample** —	
Mean housing quality	(1) Baseline	(2) With interactions
Unit & Tenant Attributes		
Building age	0.0164***	0.0159***
	(0.0030)	(0.0030)
Rent-controlled	0.0053	0.2066***
	(0.0372)	(0.0703)
Large multiunit	−0.0826**	−0.0860**
	(0.0366)	(0.0355)
Older tenant	0.2664***	0.2641***
	(0.0323)	(0.0325)
Younger tenant	−0.3396***	−0.3428***
	(0.0340)	(0.0338)
Black tenant	0.1597***	0.1604***
	(0.0381)	(0.0392)
Household size	0.0023	0.0003
	(0.0093)	(0.0093)
Married	0.0972***	0.0967***
	(0.0307)	(0.0307)
College-educated	−0.2750***	−0.2770***
	(0.0308)	(0.0306)
Household income	−0.1859***	−0.1836***
	(0.0119)	(0.0119)
Maintenance Indicators		
Adequacy rating	0.0523***	0.0533**
	(0.0428)	(0.0241)
Inside leaks	−0.0023	−0.0067
	(0.0294)	(0.0290)
Outside leaks	−0.0357	−0.0337
	(0.0319)	(0.0325)
Sewage breakdown	−0.0382	−0.0352
	(0.0758)	(0.0766)
Interaction Terms		
Post		0.1621
		(0.05787)
Treatment		−0.2925***
		(0.0791)
Owner-occupied		−0.1194***
		(0.0393)
Treatment × post		0.1204**
		(0.0614)
Owner-occupied × post		0.2352***
		(0.0596)
Treatment × owner-occupied		0.1680***
		(0.0629)
Treatment × owner-occupied × post		−0.2815***
		(0.0871)
Location fixed effects	Yes	Yes
Year fixed effects	Yes	Yes

This table presents selected results from estimation of the quality component in (3.4). The sample consists of 5,097 observations from 1985 to 2011 and includes 5 counties in total. All model specifications include year and location fixed effects. Standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

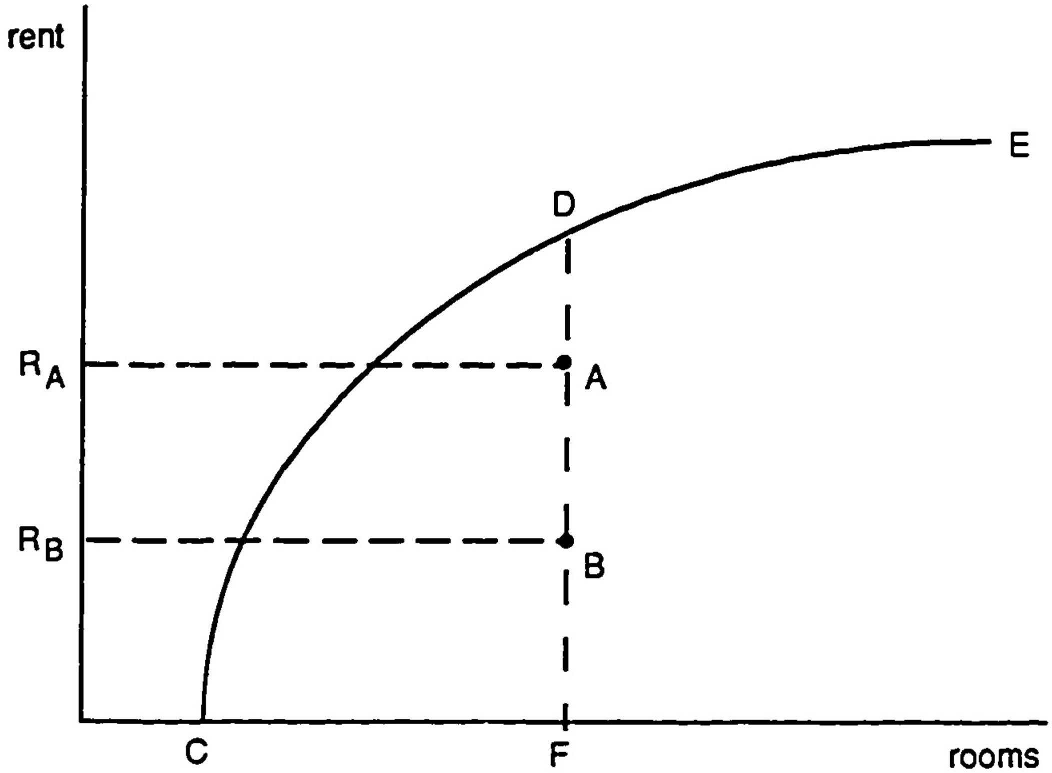


Figure 3.1: Rent Frontier

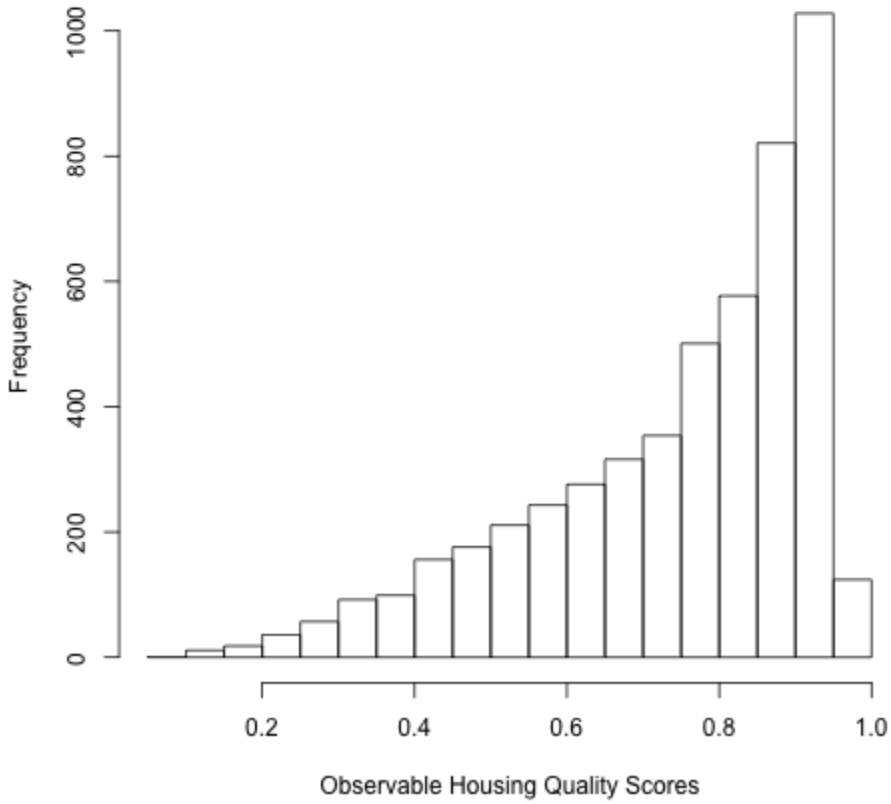


Figure 3.2: Distribution of Housing Quality Estimates

Figure 3.3: San Francisco Metropolitan Statistical Area

Appendices

Appendix A   Measure G (Proposition G)

PRINCIPAL PROVISIONS: Proposition G would change the owner move-in (OMI) provisions of the rent control law. It would

1. limit landlord evictions to one per building;
2. allow evictions for relative move-ins only if the landlord lives in or is trying to move into the same building;
3. require the landlord or relative to move in within 3 months and occupy the unit for 36 continuous months for an eviction to be legal;
4. limit an OMI eviction if the landlord has a vacant comparable unit in the building;
5. require a landlord to offer a tenant being evicted another available rental unit at a rent comparable to the original rental unit;
6. provide that a domestic partner has rights as a "spouse";
7. permanently ban the evictions of long-term senior, disabled and terminally ill tenants.

Under current law an owner in good faith may evict a tenant from a unit that the landlord or close relative will occupy. The tenant may not be evicted if there is a comparable vacant unit in the same building. The landlord or relative must live in the unit for twelve continuous months. If not, the eviction was not in good faith. Since 1983, San Francisco's Condominium Conversion Law in effect limits condominium conversions to 200 a year. It also requires relocation for tenants and requires that tenants be given first choice to buy units. However, by changing how deeds are held, "tenancies-in-common" are exempt from condominium laws. Limiting owner move-ins to one per building would eliminate the practice of using OMI evictions for tenancies-in-common in which multiple owners own a building, each living in a separate unit. Proposition G would amend Chapter 37 of the San Francisco Administrative Code, The Residential Rent Stabilization and Arbitration Ordinance (the Rent Control Ordinance). Any part of Proposition G could be changed by a vote of the Board of Supervisors without permission of the voters.

References

Afonso, F. S. (2008). Assessing and explaining the relative efficiency of local government. *Journal of Socio-Economics*, *37*, 1946–1979.

Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*, 21–37.

Anders, A., & Gearhart, R. (2018). The financial state of municipalities and the effect on housing values. *Journal of Housing Research*, *27*(1), 17–44.

Arnott, R., & Igarashi, M. (2000). Rent control, mismatch costs and search efficiency. *Regional Science and Urban Economics*, *30*(3), 249–288.

Arnott, R., & Meulbroek, L. (2018, Aug). *The stealth pension mortgage on your house.* Dow Jones Company. Retrieved from `https://www.wsj.com/articles/the-stealth-pension-mortgage-on-your-house-1533496243?mod=searchresults&page=1&pos=3`

Athanassopoulos, A. D., & Triantis, K. P. (1998). Assessing aggregate cost efficiency and the related policy implications for Greek local municipalities. *INFOR: Information Systems and Operational Research*, *36*, 66–83.

Baicker, K., Clemens, J., & Singhal, M. (2012). The rise of the states: US fiscal decentralization in the postwar period. *Journal of public Economics*, *96*, 1079–1091.

Balaguer-Coll, M. T., Prior, D., & Tortosa-Ausina, E. (2007a). On the determinants of local government performance: A two-stage nonparametric approach. *European Economic Review*, *51*, 425–451.

Balaguer-Coll, M. T., Prior, D., & Tortosa-Ausina, E. (2007b). On the determinants of local government performance: A two-stage nonparametric approach. *European Economic Review*, *51*, 425–451.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*, 1078–1092.

Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic engel curves and consumer demand. *Review of Economics and Statistics*, *79*, 527–539.

Battese, G. E. (1993). A stochastic frontier production function incorporating a model for technical inefficiency effects. *Working Papers in Econometrics and Applied Statistics*, *70*.

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*(2), 325–332.

Berry, W. D., & Lowery, D. (1984). The growing cost of government: A test of two explanations. *Social Science Quarterly*, *65*, 735–749.

Brasington, D. M., & Hite, D. (2005). Demand for environmental quality: a spatial hedonic analysis. *Regional Science and Urban Economics*, *35*(1), 57–82.

Breger, G. E. (1967). The concept and causes of urban blight. *Land Economics*, *43*(4), 369–376.

Brinkman, J., Coen-Pirani, D., & Sieg, H. (2016). *The political economy of underfunded municipal pension plans* (Tech. Rep.). National Bureau of Economic Research.

Bronshtein, G. (2017). *The consequences of the housing boom on local government debt* (Tech. Rep.).

Brueckner, J. K., & Helsley, R. W. (2011). Sprawl and blight. *Journal of Urban Economics*, *69*(2), 205–213.

Brueckner, J. K., & Rosenthal, S. S. (2009). Gentrification and neighborhood housing cycles: Will America's future downtowns be rich? *The Review of Economics and Statistics*, *91*(4), 725–743.

Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, *43*(1), 91–107.

Bruère, H., Allen, W. H., Cleveland, F. A., & Baker, S. J. (1912). Efficiency in city government. *The Annals of the American Academy of Political and Social Science*, *41*, 3–22.

Capps, K. (2014, Jun). *Peak eviction came to San Francisco back in the 1990s.* Retrieved from `http://www.citylab.com/equity/2014/06/peak-eviction-came-to-san-francisco-back-in-the-1990s/372040/`

Caudill, S. B. (1993). Estimating the costs of partial-coverage rent controls: A stochastic frontier approach. *The Review of Economics and Statistics*, 727–731.

Chalfant, J. A., & Gallant, A. R. (1985). Estimating substitution elasticities with the Fourier cost function. *Journal of Econometrics*, *28*, 205–222.

Charnes, A., W. Cooper, W., & Li, S. (1989). Using DEA to evaluate relative efficiencies in the economic performance of chinese cities. *Socio-Economic Planning Sciences*, *23*, 325-344.

Chay, K. Y., & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, *113*(2), 376–424.

Cheung, S. N. (1975). Roofs or stars: the stated intents and actual effects of a rents ordinance. *Economic Inquiry*, *13*(1), 1–21.

Cheung, S. N. (1979). Rent control and housing reconstruction: The postwar experience of prewar premises in Hong Kong. *The Journal of Law and Economics*, *22*(1), 27–53.

Chinloy, P. T. (1978). Depreciation, adverse selection and housing markets. *Journal of Urban Economics*, *5*(2), 172–187.

Cooper, R. J., & McLaren, K. R. (1996). A system of demand equations satisfying effectively global regularity conditions. *Review of Economics and Statistics*, *78*, 359–364.

Da Cruz, N. F., & Marques, R. C. (2014). Revisiting the determinants of local government performance. *Omega*, *44*, 91–103.

Daouia, A., Simar, L., & Wilson, P. W. (2017). Measuring firm performance using nonparametric quantile-type distances. *Econometric Reviews*, *36*, 156–181.

Daraio, C., Simar, L., & Wilson, P. W. (2018). Central limit theorems for conditional efficiency measures and tests of the 'separability condition' in non-parametric, two-stage models of production. *The Econometrics Journal*, *21*, 170–191. doi: 10.1111/ectj.12103

Davis, M. L., & Hayes, K. J. (1993). The demand for good government. *Review of Economics and Statistics*, *75*, 148–152.

De Borger, B., & Kerstens, K. (1996). Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches. *Regional Science and Urban Economics*, *26*, 145–170.

De Borger, B., & Kerstens, K. (2000). What is known about municipal efficiency? the Belgian case and beyond. In J. Blank (Ed.), *Public provision and performance: Contributions from efficiency and productivity measurement* (pp. 299–330). Amsterdam: Elsevier.

De Oliveira Junqueira, M. (2015). *Efficiency measurement of local public sector in international perspective: A literature survey.* (Unpublished working paper, University of São Paulo, Brazil. Presented at International Conference on Public Policy, Milan, 2015)

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor inefficiency in post offices. In M. M. P. Pestieau & H. Tulkens (Eds.), *The performance of public enterprises: Concepts and measurements* (pp. 243–267). Amsterdam: North-Holland.

Diamond, R., McQuade, T., & Qian, F. (2019). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco. *American Economic Review*, *109*(9), 3365–3394.

Doucet, B., van Kempen, R., & van Weesep, J. (2011). 'We're a rich city with poor people': municipal strategies of new-build gentrification in Rotterdam and Glasgow. *Environment and Planning A*, *43*(6), 1438–1454.

Duncombe, W., Miner, J., & Ruggiero, J. (1997). Empirical evaluation of bureaucratic models of inefficiency. *Public Choice*, *93*, 1–18.

Early, D. W. (2000). Rent control, rental housing supply, and the distribution of tenant benefits. *Journal of Urban Economics*, *48*(2), 185–204.

Early, D. W., & Olsen, E. O. (1998). Rent control and homelessness. *Regional Science and Urban Economics*, *28*(6), 797–816.

Färe, R. (1988). *Fundamentals of production theory.* Berlin: Springer-Verlag.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The measurement of efficiency of production.* Boston: Kluwer-Nijhoff Publishing.

Färe, R., & Lovell, C. A. K. (1988). Aggregation and efficiency. In W. Eichhorn (Ed.), *Measurement in economics* (pp. 639–647). Heidelberg: Physica-Verlag.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal*

Statistical Society A, *120*, 253–281.
Francke, M. K., & van de Minne, A. M. (2017). Land, structure and depreciation. *Real Estate Economics*, *45*(2), 415–451.
Galbraith, C. (1998). *Old houses never die: Assessing the effectiveness of filtering as a low-income housing policy* (Unpublished doctoral dissertation). University of Texas Austin.
García, J., Montolio, D., & Raya, J. M. (2010). Local public expenditures and housing prices. *Urban Studies*, *47*(7), 1501–1512.
Geys, B., & Moesen, W. (2009). Measuring local governmental technical (in)efficiency. *Public Performance and Management Review*, *32*, 499–513.
Glaeser, E. L. (2002). *Does rent control reduce segregation?* (Harvard Institute Research Working Paper)
Gorina, E., Joffe, M. D., & Maher, C. (2018). Using fiscal ratios to predict local fiscal distress.
Greene, W. H. (2000). Econometric Analysis 4th edition. *International edition, New Jersey: Prentice Hall*, 201–215.
Grosskopf, S., Hayes, K. J., Taylor, L. L., & Weber, W. L. (2001). On the determinants of school district efficiency: Competition and monitoring. *Journal of Urban Economics*, *49*, 453–478.
Grossman, P. J., Panayiotis, M., & Wassmer, R. W. (1999). Public sector technical inefficiency in large U.S. cities. *Journal of Urban Economics*, *46*, 278–299.
Grossman, P. J., & West, E. G. (1994). Federalism and the growth of government revisited. *Public Choice*, *79*, 19–32.
Guilkey, D. K., Lovell, C. A. K., & Sickles, R. C. (1983). A comparison of the performance of three flexible functional forms. *International Economic Review*, *24*, 591–616.
Gyourko, J., & Linneman, P. (1990). Rent controls and rental housing quality: A note on the effects of New York City's old controls. *Journal of Urban Economics*, *27*(3), 398–409.
Hájková, V., & Hájek, P. (2014). Efficiency of knowledge bases in urban population and economic growth – evidence from european cities. *Cities*, *40*, 11–22.
Harding, J., Miceli, T. J., & Sirmans, C. (2000). Do owners take better care of their housing than renters? *Real Estate Economics*, *28*(4), 663–681.
Harding, J., Rosenthal, S. S., & Sirmans, C. (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. *Journal of Urban Economics*, *61*(2), 193–217.
Hastie, T., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *55*(4), 757–779.
Hayes, K. J., & Chang, S. (1990). The relative efficiency of city manager and mayor-council forms of government. *Southern Economic Journal*, *57*, 167–177.
Hayes, K. J., Razzolini, L., & Ross, L. B. (1998). Bureaucratic choice and nonoptimal provison of public goods: Theory and evidence. *Public Choice*, *94*, 1–20.

Hayes, K. J., & Taylor, L. L. (1996). Neighborhood school characteristics: what signals quality to homebuyers? *Economic Review-Federal Reserve Bank of Dallas*, 2–9.

Hayes, K. J., & Wood, L. (1995). Utility maximizing bureaucrats: The bureaucrat's point of view. *Public Choice*, *82*, 69–83.

Henderson, J. V. (1977). *Economic theory and the cities.* Academic Press.

Henderson, J. V., & Ioannides, Y. M. (1986). Tenure choice and the demand for housing. *Economica*, 231–246.

Hilber, C. A. (2017). The economic implications of house price capitalization: a synthesis. *Real Estate Economics*, *45*(2), 301–339.

Hulten, C. R., & Wykoff, F. (1980). Economic depreciation and the taxation of structures in United States manufacturing industries: an empirical analysis. In *The measurement of capital* (pp. 83–120). University of Chicago Press.

Kiel, K. A., & Zabel, J. E. (1999). The accuracy of owner-provided house values: The 1978–1991 American Housing Survey. *Real Estate Economics*, *27*(2), 263–298.

Kneip, A., Park, B., & Simar, L. (1998). A note on the convergence of nonparametric DEA efficiency measures. *Econometric Theory*, *14*, 783–793.

Kneip, A., Simar, L., & Wilson, P. W. (2008). Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models. *Econometric Theory*, *24*, 1663–1697.

Kneip, A., Simar, L., & Wilson, P. W. (2015). When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores. *Econometric Theory*, *31*, 394–422.

Kneip, A., Simar, L., & Wilson, P. W. (2016). Testing hypotheses in nonparametric models of production. *Journal of Business and Economic Statistics*, *34*, 435–456.

Kneip, A., Simar, L., & Wilson, P. W. (2018). *Infernece in dynamic, nonparametric models of production: Central limit theorems for Malmquist indices.* (Discussion paper #2018/10, Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium)

Knight, J. R., & Sirmans, C. (1996). Depreciation, maintenance, and housing prices. *Journal of Housing Economics*, *5*(4), 369–389.

Lang, M. (1986). Measuring economic benefits from gentrification. *Journal of Urban Affairs*, *8*(4), 27–39.

Leigh, W. A. (1980). Economic depreciation of the residential housing stock of the United States, 1950-1970. *The Review of Economics and Statistics*, 200–206.

Linneman, P. (1980). Some empirical results on the nature of the hedonic price function for the urban housing market. *Journal of Urban Economics*, *8*(1), 47–68.

lo Storto, C. (2013). Evaluating technical efficiency of Italian major municipalities: a data envelopment analysis model. *Procedia-Social and Behavioral Sciences*, *81*, 346–350.

MacKay, R. C. (2014). Implicit debt capitalization in local housing prices: An example of unfunded pension liabilities. *National Tax Journal*, *67*(1), 77.
Malpezzi, S., Ozanne, L., & Thibodeau, T. G. (1987). Microeconomic estimates of housing depreciation. *Land Economics*, *63*(4), 372–385.
McAllister, P. H., & McManus, D. (1993). Resolving the scale efficiency puzzle in banking. *Journal of Banking and Finance*, *17*, 389–405.
McDonald, B. (2017). Measuring the fiscal health of muncipalities.
Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, *18*, 435–444.
Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 435–444.
Mengle, D. (1985). Logit analysis of the effect of rent control on housing quality.
Mieszkowski, P., & Zodrow, G. R. (1989). Taxation and the tiebout model: the differential effects of head taxes, taxes on land rents, and property taxes. *Journal of Economic Literature*, *27*(3), 1098–1146.
Mills, E. S., & Hamilton, B. W. (1994). *Urban Economics.* Addison-Wesley Educational Publishers, Inc.
Mitchell, K., & Onvural, N. M. (1996). Economies of scale and scope at large commercial banks: Evidence from the Fourier flexible functional form. *Journal of Money, Credit, and Banking*, *28*, 178–199.
Moon, C.-G., & Stotsky, J. G. (1993). The effect of rent control on housing quality change: a longitudinal analysis. *Journal of Political Economy*, *101*(6), 1114–1148.
Moorhouse, J. C. (1972). Optimal housing maintenance under rent control. *Southern Economic Journal*, 93–106.
Munch, J. R., & Svarer, M. (2002). Rent control and tenancy duration. *Journal of Urban Economics*, *52*(3), 542–560.
Narbón-Perpiñá, I., & De Witte, K. (2018). Local governments' efficiency: a systematic literature review—–part I. *International Transactions in Operational Research*, *25*, 431–468.
Nechyba, T. J., & Walsh, R. P. (2004). Urban sprawl. *Journal of Economic Perspectives*, *18*(4), 177–200.
Nikolov, M., & Hrovatin, N. (2013). Cost efficiency of Macedonian municipalities in service delivery: Does ethnic fragmentation matter? *Lex Localis*, *11*, 743–776.
Oates, W. E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis. *Journal of Political Economy*, *77*(6), 957–971.
Olesen, O. B., & Petersen, N. C. (2016). Stochastic data envelopment analysis—A review. *European Journal of Operational Research*, *251*, 2–21.
Olsen, E. O. (1972). An econometric analysis of rent control. *Journal of Political*

Economy, *80*(6), 1081–1100.

Pacheco, F., Sanchez, R., & Villena, M. (2014). *A longitudinal parametric approach to estimate local government efficiency.* (Unpublished working paper, Fiscalia Nacional Economica, Universidad Adolfo Ibañez, Santiago de Chile)

Palmquist, R. B. (1979). Hedonic price and depreciation indexes for residential housing: A comment. *Journal of Urban Economics*, *6*(2), 267–271.

Park, B. U., Simar, L., & Weiner, C. (2000). FDH efficiency scores from a stochastic point of view. *Econometric Theory*, *16*, 855–877.

Pope, D. G., & Pope, J. C. (2012). Crime and property values: Evidence from the 1990s crime drop. *Regional Science and Urban Economics*, *42*(1-2), 177–188.

Radulović, B., & Dragutinović, S. (2015). Efficiency of local self-governments in Serbia: an SFA approach. *Industrija*, *43*, 123–142.

Randolph, W. C. (1988). Estimation of housing depreciation: Short-term quality change and long-term vintage effects. *Journal of Urban Economics*, *23*(2), 162–178.

Revelli, F. (2009). Spend more, get more? an inquiry into english local government performance. *Oxford Economic Papers*, *62*, 185–207.

Robins, P. K., & West, R. W. (1977). Measurement errors in the estimation of home value. *Journal of the American Statistical Association*, *72*(358), 290–294.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, *56*, 931–954.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55.

Schmidt, P., & Lin, T.-F. (1984). Simple tests of alternative specifications in stochastic frontier models. *Journal of Econometrics*, *24*(3), 349–361.

Shephard, R. W. (1970). *Theory of cost and production functions.* Princeton: Princeton University Press.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of productive efficiency. *Journal of Econometrics*, *136*(1), 31–64.

Simar, L., & Wilson, P. W. (2011). Two-Stage DEA: Caveat emptor. *Journal of Productivity Analysis*, *36*, 205–218.

Simar, L., & Wilson, P. W. (2013). Estimation and inference in nonparametric frontier models: Recent developments and perspectives. *Foundations and Trends in Econometrics*, *5*, 183–337.

Simar, L., & Wilson, P. W. (2015). Statistical approaches for non-parametric frontier models: A guided tour. *International Statistical Review*, *83*, 77–110.

Simar, L., & Wilson, P. W. (2019). *Hypothesis testing in nonparametric models of production using multiple sample splits.* (Discussion paper 2019/19, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium)

Skidmore, M., & Scorsone, E. (2011). Causes and consequences of fiscal stress in Michigan cities. *Regional Science and Urban Economics*, *41*(4), 360–371.

Štastná, L., & Gregor, M. (2015). Public sector efficiency in transition and beyond: evidence from czech local governments. *Applied Economics*, *47*, 680–699.

Tang, K.-L. (1997). Efficiency in the private sector: A critical review of empirical evidence from public services. *International Review of Administrative Sciences*, *63*, 459–474.

Thompson, P. N. (2016). School district and housing price responses to fiscal stress labels: Evidence from Ohio. *Journal of Urban Economics*, *94*, 54–72.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64*(5), 416–424.

Tsoodle, L. J., & Turner, T. M. (2008). Property taxes and residential rents. *Real Estate Economics*, *36*(1), 63–80.

Walters, C. R. (2009). Do subsidized housing units depreciate faster than unsubsidized ones? *Journal of Housing Economics*, *18*(1), 49–58.

Wan, A. T., Xie, S., & Zhou, Y. (2017). A varying coefficient approach to estimating hedonic housing price functions and their quantiles. *Journal of Applied Statistics*, *44*(11), 1979–1999.

Wang, X., Dennis, L., & Tu, Y. S. (2007). Measuring financial condition: A study of US states. *Public Budgeting & Finance*, *27*(2), 1–21.

Wheelock, D. C., & Wilson, P. W. (2001). New evidence on returns to scale and product mix among U.S. commercial banks. *Journal of Monetary Economics*, *47*, 653–674.

Wheelock, D. C., & Wilson, P. W. (2008). Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations. *Journal of Econometrics*, *145*, 209–225.

Wheelock, D. C., & Wilson, P. W. (2012). Do large banks have lower costs? new estimates of returns to scale for U.S. banks. *Journal of Money, Credit, and Banking*, *44*, 171–199.

Wheelock, D. C., & Wilson, P. W. (2018). The evolution of scale-economies in U.S. banking. *Journal of Applied Econometrics*, *33*, 16–28.

Wilhelmsson, M. (2008). House price depreciation rates and level of maintenance. *Journal of Housing Economics*, *17*(1), 88–101.

Wilson, P. W. (2011). Asymptotic properties of some non-parametric hyperbolic efficiency estimators. In I. Van Keilegom & P. W. Wilson (Eds.), *Exploring research frontiers in contemporary statistics and econometrics* (pp. 115–150). Berlin: Springer-Verlag.

Wilson, P. W. (2018). Dimension reduction in nonparametric models of production. *European Journal of Operational Research*, *267*, 349–367.

Wilson, P. W. (2019). U.S. banking in the post-crisis era: New results from new methods.
(forthcoming)

Wilson, P. W., & Carey, K. (2004). Nonparametric analysis of returns to scale and product mix among US hospitals. *Journal of Applied Econometrics*, *19*, 505–524.

Worthington, A. C. (2000). Cost efficiency in Australian local government: A comparative analysis of mathematical programming and econometrical approaches. *Financial Accountability and Management*, *16*, 201–223.

Yoshida, J. (2016). Structure depreciation and the production of real estate services. *Available at SSRN 2734555*.

Zahirovic-Herbert, V., & Chatterjee, S. (2012). Historic preservation and residential property values: Evidence from quantile regression. *Urban Studies*, *49*(2), 369–382.

Zhang, L. (2016). Flood hazards impact on neighborhood house prices: A spatial quantile regression analysis. *Regional Science and Urban Economics*, *60*, 12–19.

Zietz, J., Zietz, E. N., & Sirmans, G. S. (2008). Determinants of house prices: A quantile regression approach. *The Journal of Real Estate Finance and Economics*, *37*(4), 317–333.